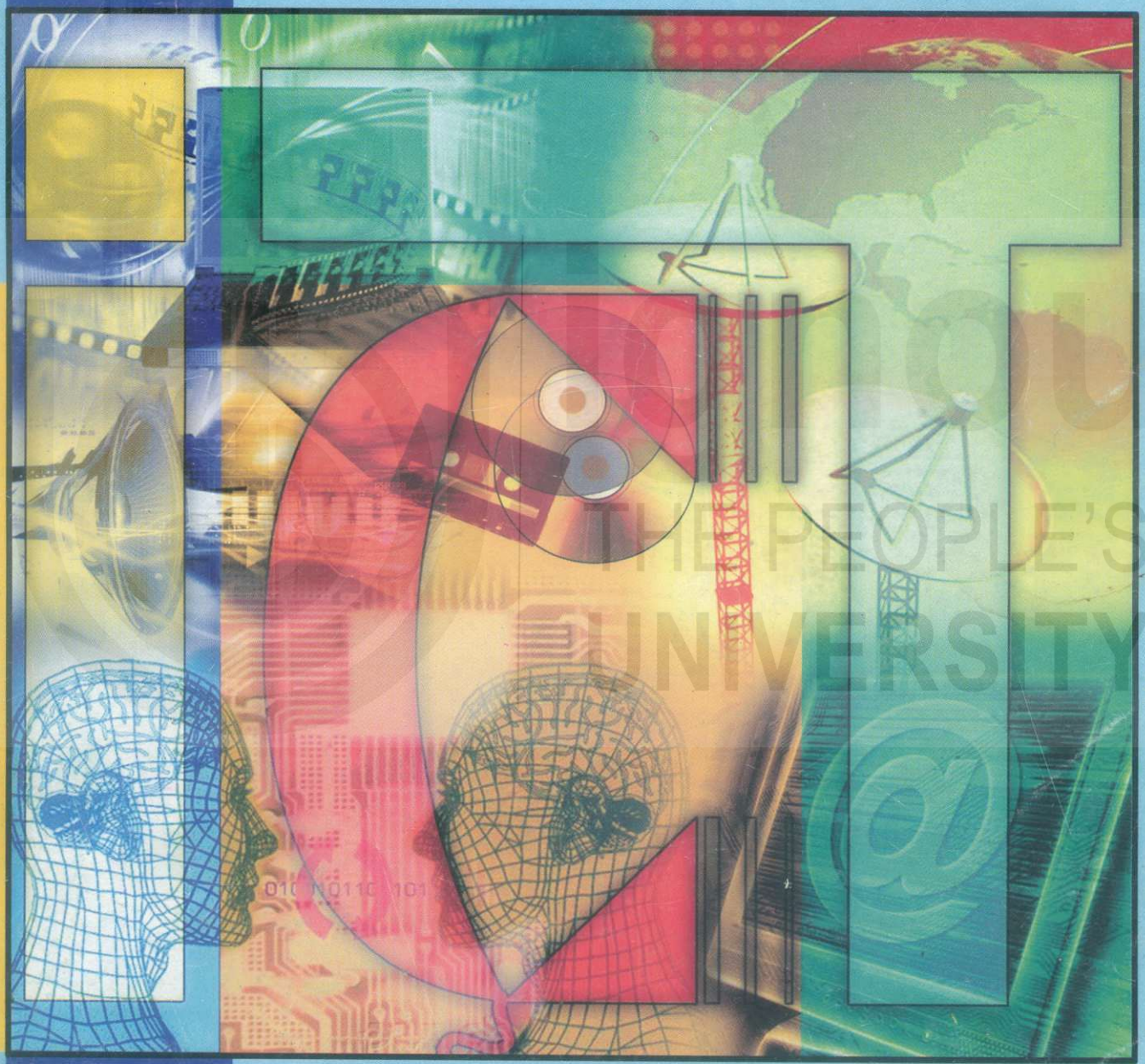
A·S

Indira Gandhi National Open University
School of Social Sciences

MLII - 103

Master's Degree in Library and Information Science (MLIS)



FUNDAMENTALS OF INFORMATION COMMUNICATION TECHNOLOGIES

"शिक्षा मानव को बन्धनों से मुक्त करती है और आज के युग में तो यह लोकतंत्र की भावना का आधार भी है। जन्म तथा अन्य कारणों से उत्पन्न जाति एवं वर्गगत विषमताओं को दूर करते हुए मनुष्य को इन सबसे ऊपर उठाती है।"

— इन्दिरा गांधी

ignou
THE PEOPLE'S
UNIVERSITY

"Education is a liberating force, and in our age it is also a democratising force, cutting across the barriers of caste and class, smoothing out inequalities imposed by birth and other circumstances."

– Indira Gandhi

Indira Gandhi
National Open University
School of Social Sciences

MLII-103

Fundamentals of Information Communication Technologies

BLOCK 1: COMPUTER TECHNOLOGY

UNIT 1 : Architecture and Technology 9

UNIT 2 : Software 36

UNIT 3 : Programming 72

BLOCK 2: COMMUNICATION TECHNOLOGY

UNIT 4 : Fundamentals of Communication Technology 89

UNIT 5 : Networks and Networking 108

UNIT 6 : Data Networks 129

UNIT 7 : Convergence Technology and Applications 155

BLOCK 3: MEDIA AND PUBLISHING TECHNOLOGY

UNIT 8 : Print and Non-Print Media 181

UNIT 9 : Printing and Publishing Technology 222

UNIT 10 : E-Publishing and Network Publishing 247

UNIT 11 : Overview of Reprography and Micrography 285

BLOCK 4: RESOURCE SHARING NETWORKS

UNIT 12 : Library and Information Networks 319

UNIT 13 : Bibliographic Utility Networks 356

UNIT 14 : Library and Information Network in India 379

ignou THE PEOPLE'S UNIVERSITY

Programme Design Committee (Original) (1992)

Prof. Pandav Nayak (*Chairman*)
Prof. (Ms.) A.K. Anand
Prof. J.C. Binwal
Prof. M.A. Gopinath
Prof. B. Guha
Prof. S.R. Gunjal
Prof. S.G. Mahajan
Mr. N.M. Malwad
Dr. S.S. Murthy
Prof. K.S. Raghavan
Prof. T.N. Rajan
Prof. A.P. Srivastava
Prof. T. Viswanathan
Dr. (Ms.) Neela Jagannathan *(Special Invitee)*
Dr. R. Satyanarayana (*Convener*)
Internal Members
Dr. Uma Kanjilal
Ms. Neena Talwar Kanungo

Course Preparation Team (1992)

Mr. J.M. Bhardwaj
Ms. S. Kar
Mr. H.K. Kaul
Dr. P.S.G. Kumar
Mr. N. M. Malwad
Dr. K.C. Panda
Mr. M.K. Raghvendra Rao
Mr. V. Ramachandran
Dr. R. Satyanarayana
Prof. T. Viswanathan
Course Editor: Prof. M.A. Gopinath

Programme (Curriculum) Revision Committee (2003)

Prof. S.B. Ghosh (*Chairman*)
Faculty of Library & Information Science,
School of Social Science,
IGNOU, New Delhi

Dr. Jagdish Arora
Librarian
Indian Institute of Technology
Mumbai

Prof. S.R. Ganpule
Dept. of Library & Information Science,
Yaswant Rao Chavan Maharashra Open University,
Nasik

Prof. B. Guha
Professor (Retd.)
Dept. of Lib. & Inf. Sc.
Banaras Hindu University,
Banaras, U.P.

Prof. Uma Kanjilal
Faculty of Lib. & Inf. Sc.
IGNOU, New Delhi

Dr. Neena Talwar Kanungo
Faculty of Lib. & Inf. Sc.
IGNOU, New Delhi

Prof. C.R. Karisiddappa
Head, Dept. of Library & Information Science,
Karnataka University,
Dharwar

Dr. A. Lahiri
Senior Advisor,
Dept. of Scientific & Industrial Research
New Delhi

Prof. P.B. Mangla
Dept. of Lib. & Inf. Sc. (Retd)
University of Delhi
Delhi

Prof. A. Neelameghan
UNESCO Expert (Retd.)
Sarada Ranganathan Endowment for Library & Information Science
Bangalore

Prof. T.N. Rajan
Chief Coordinator (Retd.)
INSDOC, New Delhi

Dr. Pravakar Rath
Faculty of Lib. & Inf. Sc.
IGNOU, New Delhi

Prof. J. Sarkhel
Head, Dept. of Library & Information Science, Vidyasagar University,
Midnapur W.B.

Prof. R. Satyanarayana
Faculty of Lib. & Inf. Sc. (Retd.)
IGNOU, New Delhi

Prof. B.K. Sen
Head, (Retd.),
Education Division, INSDOC,
New Delhi

Dr. Mahinder Singh
Director
DESIDOC, Delhi

Dr. N. Vijayaditya
Director General
National Informatics Centre
New Delhi

Prof. T. Viswanathan
Ex-Director, INSDOC, New Delhi

Dr. (Ms) Neela Jagannathan (*Special Invitee*)
IGNOU, New Delhi

Programme Coordinators: Prof. S.B. Ghosh and Prof. Uma Kanjilal

Course Coordinator: Prof. S.B. Ghosh **Course Editor**: Dr.V.S.P. Srivastava

Course Preparation Team

Unit	Contributor	Academic Assistance
1	Prof. R. Satyanarayana	Mr. Anup Kumar Das, Consultant
2	Dr. P.K. Panigrahi	
3	Mr. G. Mahesh	
4	Prof. T. Vishwanathan	
5	Prof. T. Vishwanathan	
6	Prof. T. Vishwanathan	
7	Prof. T. Vishwanathan	
8	Prof. A.Y. Asundi	
9	Mr. V. Ramachandran	
10	Dr. P.K. Panigrahi	
11	Mr. M.K. Raghvendra Rao	
12	Dr. J. Arora	
13	Dr. J. Arora	
14	Dr. Aditya Tripathi	

Secretarial Assistance Shri S. Swaminathan

Material Production

Mr. Jitender Sethi
School of Social Sciences
IGNOU, New Delhi

Mr. S. S. Venkatachalam and
Mr. Manjit Singh
School of Social Sciences
IGNOU, New Delhi

September, 2005

© Indira Gandhi National Open University, 2005

ISBN-81-266-2019-6

All rights reserved. No part of this work may be reproduced in any form, by mimeography or any other means, without permission in writing from the Indira Gandhi National Open University.

Authors are responsible for the academic content of course as far as the copyright issues are concerned.

Further information on the Indira Gandhi National Open University courses may be obtained from the University's Office at Maidan Garhi, New Delhi-110 068

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi, by Director, School of Social Sciences.

"Paper Used: Agrobased Environment Friendly"

Laser Typeset by : Rajshree Computers, 5A/177, W.E.A., Karol Bagh, New Delhi-110 005

Printed at M/s Dee Kay Printers, 5/37A, Kirti Nagar Indl. Area, New Delhi-110015.

FUNDAMENTALS OF INFORMATION COMMUNICATION TECHNOLOGIES

Introduction

The course entitled as 'Fundamentals of Information Communication Technologies' has the objective to let the students know about the developments and exploitation of the ICT in various fields with particular reference to its applications in the field of Library and Information Sciences.

Information Technology has undoubtedly been playing significant role in every facet of life. IT, clubbed with the Communication Technology has a great potential in providing support to one and all living even in the most remote area worldwide. In a span of just 50 years we see at least five generations of computers made of electronic components. Developments and advancements in computer hardware, both in terms of speed and storage capacity, have given us tremendous support towards information management. Almost every desk today in various offices uses IT for some or the other businesses, libraries and information centres being no exception.

Data processing, information generation, its storage, and dissemination depend on many other constituents. Starting with computer hardware and the requisite software it gets combined with network infrastructure and various communication technologies to create the ICT environment ultimately. Internet facility using various ways of connectivity like leased line through PSTN, point-to-point fiber optic connectivity, and VSAT— a greatest tool available today for establishing connectivity around the globe. Through IP address any machine can be identified over the worldwide network. Library applications for database, information storage and sharing are very much benefited from this technology like any other field of applications.

For the library and information professionals, it is necessary that they should know the basics of ICT and their features so that they are in a convenient position to apply ICT according to the various requirements of libraries and information centres.

This course elaborates on fundamentals of ICT — tools and media, starting with their genesis. There are fourteen Units giving details about the various constituents. **Units 1 to 3** give details about computer hardware and software aspects. **Unit 4** gives details about various means of communication starting from its genesis to the present scenario. **Units 5 and 6** discuss about network and networking concepts with details about various data networks available. **Units 8 to 11** provide details about the various output media and technology starting with paper, microforms, audiovisual and electronic media. Multimedia and its applications have also been discussed. The printing and publishing technology is undergoing tremendous change due to the development of technology. Developments in this field including electronic and network publishing has been discussed in **Units 9 and 10. Unit 11** discusses the reprography and micrography technology. **Units 12 to 14** give details about Library and Information Networks worldwide and in India covering also the details about bibliographic utility networks. This course will provide you enough insight in the development of ICT and its features to act as a take-off point to understand the applications of ICT in the field of Library and Information Work discussed in the next course- MLII-104.

BLOCK 1

COMPUTER TECHNOLOGY

UNIT 1 ARCHITECTURE AND TECHNOLOGY

Structure

1.0 Objectives

1.1 Introduction

1.2 Computer Generations

1.3 Processor Technology

- 1.3.1 Microprocessor Technology
- 1.3.2 Computer Architecture: Historical Perspective
- 1.3.3 Basic Computer Architecture

1.4 Storage Technology

- 1.4.1 Main Memory
- 1.4.2 Magnetic Storage
- 1.4.3 Optical Storage

1.5 Input / Output Devices

- 1.5.1 Input Devices
- 1.5.2 Output Devices

1.6 Summary

1.7 Answers to Self Check Exercises

1.8 Keywords

1.9 References and Further Reading

1.0 OBJECTIVES

Information Communication Technologies (ICT) has resolutionised the occurs, storage, retrieval and dissemination of information. This has been possible due to the developments in components of ICT of which Computer Technology is one. In this Unit we will be discussing about the fundamentals of Computer Technology.

After reading this unit, you will be able to:

- know what constitutes information technology;
- describe the significant features of different generations of computers;
- understand the concept of computer architecture;
- know the developments in processor and storage technologies; and
- comprehend the range of different input/output devices available today.

1.1 INTRODUCTION

Technology has been defined as "systematic knowledge and action, usually of industrial process but applicable to any recurrent activity". In providing tools and techniques for action, technology at once adds to and draws from a knowledge base in which *theory and practice interact and compact.* At its most general level, technology may be regarded as definable, specifiable way of doing any thing. In other words, we may say that technology is a *codified communicable procedure for solving problems.*

Information Technology (IT) is a term that has only been heard since late 1970s. Basically, the term refers to the technologies and applications developed through data processing and storage powers of computers with the distance transmission capabilities of telecommunication technologies which previously existed as different industries with separate functions. Computers were first produced and used to manipulate and store data rather than for distance communication, while telecommunications were mainly concerned with transmission of speech (via telephones) or in audio and video form which were basically mass media broadcasts. Over the last thirty years however, these separate technologies and their functions have been coming together. This has been made possible by new technical developments, particularly in the miniaturization of electronic circuits. It is *this recent convergence of computing and telecommunications based on microelectronics,* together with its various applications in the modern world *that people usually mean by Information Technology (IT).* In fact the foundation of technological revolution that is characterised today, as the architect of the "Information Age" can be found in the relationship among information, the computer and communications. The application of microelectronics to these components has generated a convergence among all the three that today and in future will continue to promote the societal transformation. With this background knowledge of IT, let us now turn our attention to the developments in computer technology.

1.2 COMPUTER GENERATIONS

The study of computer technology begins with the invention of computer and its evolution into a powerful machine. All this happened over a period of time and is based on many technological developments. It is stated that modern computer systems have gone through five generations of evolution. Let us try to understand as to how all this happened.

From early 1940s till the present, computer scientists have been able to identify some clear-cut stages in the development of computer technology. With each stage of development, radical break throughs in electronics occurred, with the result that computers based on older technology of electronics have been replaced by a newer form of machines. As a consequence, computer systems have gone through five generations of evolution.

The first generation machines seen in the late 1940s and early 1950s used vacuum tubes for the central processor construction and magnetic core devices for main memory. The interface between man and the machine (computer) was through machine language consisting of *ones* and *zeros* (1s and 0s). At a later point in time, *mnemonic symbols* were used to replace *ones* and *zeros* and a software package known as *"assembler"* was used to translate the *mnemonic symbols* into the machine language. The first generation of computers were giant, special purpose, consumed high power and were prone to failures because of the low reliability of the vacuum tubes. In addition, they were slow in operation and possessed limited computing capacity. They were used for complex numerical calculations such as finding out the trajectories of missiles etc. Since large amount of heat was generated during operation, intense air conditioning was essential for their maintenance.

The *second generation* of computers came into being in the early 1960s with the advent of *transistor* technology. Transistors are semiconductor devices that are functionally equivalent to vacuum tubes but are smaller in size. Their advantage as low power consumption, low cost and more reliable replacement for vacuum tubes was quickly recognised. As a result they replaced vacuum tubes in large number. Computers built using transistors were compact in size, and occupied less space. Because of the compactness of the machine, the term minicomputer came into use. High-level programming languages like FORTRAN were developed and used as human-machine interface languages. Preliminary concepts of operating systems for resource management were also introduced in the design and building of the second-generation machines. It may be mentioned that use of transistors fastened the processing speed and reduced the computational time from *milliseconds* to *microseconds*. Other advantages associated with the second-generation machines were less hardware failures and better portability.

Because of these advantages found in second generation, use of computers came into existence for commercial purpose.

New technological discoveries rapidly found their way into the new models of computers. In 1962, a team of scientists successfully fabricated a number of transistors on a *single semiconductor wafer* (silicon wafer). They were able to interconnect these transistors to realise certain 'logic functions' that could be used as building blocks for computers. Such devices were known as *integrated circuits* (ICs), as many transistors were integrated on a *single wafer.* As a result of this development, *third generation* of computers came into being in 1970s. This technology further reduced the size and power consumption of computers and vastly improved their reliability. It so happened that during this period many high level languages like ALGOL, COBOL, etc. were developed for providing human machine interface. It is at this juncture, that operating systems were developed to provide 'multiprogramming' and 'time-sharing' features. Invent of multi-programming concept increased the overall performance of processor, the computer brain. These developments, as such, contributed to the introduction of modes of operation enabling the sharing of high capability available. One such mode was known as batch processing in which problems were prepared and held ready for computation on a relatively inexpensive storage medium such as magnetic drums, magnetic 'disk packs', or magnetic tapes. This mode of operation is used even today for processing large volume of data. Another mode of operation for fast, powerful machines was called *time-sharing.* In time-sharing, the computer processes many waiting jobs in such rapid succession that each job progresses quickly as if other jobs did not exist, thus, keeping each customer satisfied. The 'time-sharing' feature of the third generation machines enabled them to be connected to networks.

Ever since the advent of 'ICs', continuous research has been going on to integrate more and more devices on a 'single wafer' of fixed area. This led to the evolution of 'small scale integration' (SSI), 'medium scale integration' (MSI), 'large scale integration' (LSI), and 'very large scale integration' (VLSI). The present technological level is 'VLSI', where in about '100,000 to a few million' devices are integrated in one 'single wafer'. In fact, the development of VLSI resulted in the emergence of what is known as *microprocessor,* which is a *whole central processing unit* (CPU) on a 'single chip'. Computers built using microprocessors are known as microcomputers. As a matter of fact, all the personal computers are known as *fourth generation machines.* The enormous success of personal computers (PCs) coupled with the advances in microprocessor technology resulted in the emergence of two major manufacturers of microprocessors namely: Motorola Inc and Intel Corporation. It may be noticed that Motorola microprocessors are numbered as 68000, 68020, 68030, 68040 and so on, on the other hand, the Intel microprocessors are numbered 8080, 8085, 8086, 80186, 80286, 80386 and 80486 and so on. Further generation of chip from Intel are known as *Pentium I/II/III/IV* with *Celeron* and *MMX* in between and *Itanium* as the latest.

It may be stated that IBM adopted Intel series of microprocessors for their personal computers. This is the reason why we always talk of 286AT, 386AT and 486AT when we consider IBM compatible personal computers. PC / XTs were designed with 8086 processor. The acronym 'XT' was used for extended technology while 'AT' for advanced technology. Along with the development of desktop microcomputers, a trend towards the building of a very large capacity computer emerged. These computers were known as *super computers.* The super computers are built using parallel architectures wherein a number of processors are clubbed together to execute jobs in parallel. These machines are both powerful and costly and form part of fourth generation machines. During this period, computer field as a whole has experienced tremendous growth. The integration of computer and telecommunications technologies has resulted in computer networking and a number of applications such as "electronic mail" and "electronic publishing" have become popular. No doubt, that the first four generations of computers were basically built to handle numeric and data processing applications and these certainly enabled to achieve that primary purpose.

At this juncture, it must be noted that the transition from one computer generation to the next takes some years of research effort and 'path breaking' advances made during the years, thus, changing the face of existing technologies. It is also difficult to be very precise about the end of one generation and the beginning of the next generation. Although, the fourth generation of computers are used in 1990s, some of these computers and computer systems have some properties which might be considered as the characteristics of fifth generation computers.

The most noticeable characteristic of the fifth generation computers is the ability to apply previously gained knowledge, to draw conclusions and then execute a task. In other words, the computer will, in short, stimulate the human ability to reason. That is to say, that computers will have to be able to classify information, search large databases rapidly, plan, apply the rules which humans regard as clear thinking, make decisions and learn from their mistakes. Input devices for the fifth generation computers could also include speech and visual recognition.

Some of the technological advances which enabled the development of fifth generation computers include:

- Parallel processing - i.e., many processors are grouped together to function as *one large group processor;*
- Super Conductors - super conductor is *a conductor through which electricity can travel without any resistance* resulting in faster transfer of information between the components of a computer;

Today's computers (those of fourth generation) already exhibit some of the characteristics of fifth generation computers. For example:

- Expert systems used by doctors to reach diagnosis following the logical steps of problem solving;
- Speech recognition systems capable of recognising dictation and entering the text into a Word processor are presently available, though these need tbe perfected.

Some of the major applications envisaged of fifth generation computers are:

- intelligent robots that could 'see' their environment (visual input e.g. a video camera) and could be programmed to carry out certain tasks without *step by step instructions.* The robot should be *able to decide for itself* how the task should be accomplished, based on observations it made of its environment. The robot ASIMO is an example of such intelligent robots;
- intelligent systems that could *control the route of a missile* and defence systems that could fend off attacks;
- word processors that could be freely controlled by means of speech recognition;
- programs that could *translate documents from one language to another;*
- pattern recognition, knowledge processing, etc.

It is stated that fifth generation computers will involve major deviations from the *know how* used for the first four. New operation of symbolic manipulation and symbolic reasoning will be required for the shift to *knowledge processing.* It appears that the conventional *Von Neumann Architecture* will be discarded in favour of *a revolutionary parallel architecture.* This will call for new *low level programming languages, methods of memory management* and *processor organisation techniques,* etc. The main features of the fifth generation computers include *inference engines, knowledge bases, natural language interface, photonic devices,* and *computing in the optical domain.* The fifth generation computers, though were expected to be in the market in 1990s, *are yet to become communicably available.* It may be pertinent to mention here that further

advancements are taking place in the fourth generation computers making them more and more powerful with added characteristic features.

In concluding this section on computer generations, it must be pointed out that beyond the third generation, there is less general agreement on defining generations of computers. There have been a fourth and fifth generation, based on advances in integrated circuit technology. With the introduction of large-scale integration (LSI), more than one thousand (1000) components can be placed on a single integrated chip. Very large scale integration (VLSI) achieved more than 10000 components per chip, and current VLSI chips can contain more than 100,000 components. With the advancement in IC miniaturisation both the processor speed and memory enlargement have very significantly been addressed and advanced. Today's personal computer is much more faster and capable than earlier mini and mainframe computers. Also, because of achievements in ICs size and integration, size of today's computer has drastically and tremendously reduced.

With the rapid pace of technology, the high rate of introduction of new products and the importance of software and communications as well as hardware, the classification by generation becomes *less clear* and *less meaningful.* It could be said that the commercial application of new developments resulted in a major change in the early 1970s and that the results of these changes are still being worked out.

Self Check Exercises

1) What is meant by Information Technology?

2) Identify the chronology of different generations of computers and mention the special features of each.

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

1.3 PROCESSOR TECHNOLOGY

The pace of technological change is revolutionary. The convergence among enabling technologies has progressed to such an extent that even for the purpose of exposition, it is difficult to separate out one technology from another. Processor technology appears to be no exception to the rule.

1.3.1 Microprocessor Technology

It may be mentioned that microchip remains central to the progress in information and communication technologies in general and that of computer technology in particular. Silicon continues to provide the basic raw material both for logic chips and memory chips, known today as Dynamic Random Access Memory Chips (DRAMC).

The first application of integrated circuit technology to computers was construction of processor. It was also found that the same technology could be used to construct memories.

Just as the density of elements on memory chips continued to rise, so also the density of elements on processor chips went on rising, so that fewer and fewer chips were needed *to construct a single computer processor.* This process led to the evolution of *a microprocessor.* The credit of developing the first *microprocessor* goes to Intel. The 4004 was the first chip to contain all the components of a CPU on a single chip and was named *microprocessor.* It marked the beginning of a continuing evolution of microprocessor capability and power.

This evolution can be seen most easily in the number of bits (zeroes and ones) that the processor deals with at a time. There is no clear cut measure of this, but perhaps the best measure is the data bus width: (i.e.) the number of bits of data that can be brought into or sent out of the processor at a time. Another measure is the number of bits in the accumulator or in the set of general purpose register which are temporary storage for use by processor. Often, these measures coincide, but not always. For example, a number of microprocessors were developed that could operate on 16 bit numbers in registers but can only read and write 8 bits at a time. By the end of 1970s, powerful general purpose 16 bit microprocessors appeared. One of these was *8086.* The next step in this trend occurred in 1981, when both Bell Labs and Hewlett Packard (HP) developed 32 bit single chip microprocessors. Intel introduced its own 32-bit microprocessor 80386 in 1985. Tables a, b, c and d present the evolution of Intel Microprocessors.

Table1.1: Evolution of Intel Microprocessors

(a) 1970s Processors

	4004	8008	8080	8086	8088
Introduced (mm/dd/yy)	11/15/71	4/1/72	4/1/74	6/8/78	6/1/79
CPU Clock Speeds	108 KHz	108 KHz	2MHz	5MHz, 8MHz, 10 MHZ	5MHz, 8MHz
Bus Width	4 bits	8 bits	8 bits	16 bits	8 bits
Number of Transistors (microns)	2300 (10)	3500	6000 (6)	29000 (3)	29000 (3)
Addressable Memory	640 bytes	16 Kbytes	64 Kbytes	1 MB	1MB
Virtual Memory	–	–	–	–	–

(b) 1980s Processors

	80286	386TM DX	386 TM SX	486 TM DX CPU
Introduced (mm/dd/yy)	2/1/82	10/17/85	6/16/88	4/10/89
CPU Clock speeds	6MHZ-12.5MHz	16MHz-33MHz	16MHz-33MHz	25MHz-50MHz
Bus width	16 bits	32 bits	16 bits	32 bits
Number of Transistors (microns)	134000 (1.5)	275000 (1)	275000 (1)	1.2 million (0.8-1)
Addressable Memory	16 megabytes	4 gigabytes	4 gigabytes	4 gigabytes
Virtual Memory	1 gigabyte	64 terabytes	64 terabytes	64 terabytes

(c) 1990s Processors

	486TM SX	Pentium	Pentium	Pentium II
Introduced (mm/dd/yy)	4/22/91	3/22/93	11/01/95	5/07/97
CPU Clock Speeds	16 MHz-133 MHz	60 MHz-166 MHz	150 MHz-200 MHz	200MHz-300 MHz
Bus Width	32 bits	32 bits	64 bits	64 bits
Number of Transistors (microns)	1.185 million (1)	3.1 million (0.8)	5.5 million (0.6)	7.5 million (0.35)
Addressable Memory	4 gigabytes	4 gigabytes	64 gigabytes	64 gigabytes
Virtual Memory	64 terabytes	64 terabytes	64 terabytes	64 terabytes

(d) Recent Processors

	Pentium III	Pentium IV
Introduced	2/26/99	11/2000
CPU Clock Speeds	450-660 MHz	1.3-1.8 GHz
Bus Width	64 bits	64 bits
Number of Transistors (microns)	95 million (0.25)	42 million
Addressable Memory	64 gigabytes	64 gigabytes
Virtual Memory	64 terabytes	64 terabytes

Source: Intel Corp. http://www.intel.com/intel/museum/25anniv/hot/tspecs.htm

Two major categories of microprocessor technology are presently being identified. They are: traditional CISC (Complex Instruction Set Computing) and more recent microprocessors RISC (Reduced Instruction Set Computing). The more instructions contained in a central processing unit (CPU) instruction set and the more complicated those instructions, the more *clock cycles* the CPU needs to complete an instruction search. Traditional CISC technology has complex sets of between 150 and 300 instructions and processes *only one instruction at a time.* The higher performance RISC chips, on the other hand, contain *70 to 80 of only the most frequently used instructions,* and are capable of processing a new instruction *on each clock cycle* (Michelle, 1991).

RISC uses smart compiler technology, which has the capability to translate commands from software language to machine language and send them to CPU, but does so, in the form of strings of most frequently used instructions broken down into pieces for storing in a reserved memory *cache.* RISC also employs *pipelining and multiple parallel processing to avoid internal bottlenecks* and *to speed up processing.*

Developments in a very long word instruction (VLWI) technology are expected to be more efficient for parallel processing purposes and this technology has potential, when fully developed, to replace RISC technology. As far as the chips themselves, gallium arsenide retains its potential for speeds up to *five or six times* those of silicon chips. New

silicon products which are *8 to 16 times* as powerful as latest version of Intel Pentium chip are expected to be in the market soon. It is also claimed that the data on CPUs will be in highly compressed form, based on specialized digital signal processors (DSPs) using data specific forms of compression (i.e.) one for audio, and one for video and so on. DSPs are fast, single function co-processor chips that relieve the CPU of many of the load requirements that accompany audio and video computing. They also eliminate one of the problems involved in putting advanced functionality in hardware because they can be reprogrammed as standards change and features grow.

It may be pointed out that this continuing technological revolution has enabled the development of applications of great importance, complexity and power. For example desktop applications that require great power of today's micro processor-based systems include:

- Image processing
- Speech recognition
- Video conferencing
- Multimedia authoring, etc.

Workstation systems which support these applications are now available in the market.

1.3.2 Computer Architecture: Historical Perspective

The term computer architecture is one of the basic concepts in computer science. There appears to be no general agreement as regards to its definition. Therefore, many interpretations could be found in the literature.

In describing computers, a distinction is often made between the two concepts: computer architecture and computer organization. Even though it is difficult to give precise definitions for these terms, a consensus exists about the general areas covered by each. (Vranesic, Z., and Thurber, K., 1980).

'Computer architecture' refers to those attributes of a system which are visible or, stated in other words, those attributes that have a direct impact on the logical execution of a program. On the other hand, 'computer organization' refers to the operational units and their interconnections that *realise* the architectural specifications. Examples of architectural attributes include the 'instruction set', the number of 'bits' used to represent various data types (numbers, characters), Input/Output (I/O) mechanisms, and techniques for addressing memory. Organisational attributes include those hardware details transparent better to the programmer, such as control signals; interfaces between the computer and the peripherals, and the memory technology used.

The distinction between architecture and organisation has been an important one. It is because many computer manufacturers offer a family of computer models, all with the same architecture but with differences in organisation. As a result, the different models in the family have different price and performance characteristics. In addition *a particular architecture* may span many years and encompass a number of different computer models, its organisation changing with changing technology.

In microcomputers, the relation between architecture and organisation is very close. Changes in technology not only influence organisation but also result in the introduction of more powerful and more complex architectures.

The term computer architecture was first coined in 1964 by the 'architects' of IBM-360 System (Amdahl, *et al).* They interpreted the term architecture as "the structure of a computer that machine language programmer must understand to write a correct programme for the machine". It comprises, as such, the definition of registers and memory as well as the instruction set, instruction formats, addressing models and the actual coding of the instructions excluding implementation and realisation. By implementation the authors

meant the actual hardware structure, and by realisation, the logic technology packaging and interconnections.

A number of other authors have also attempted to interpret the concept in their own way.

John Von Neumann, who is considered to be the father of modern computers, gave us two part architecture for modern computer systems. This architecture consists of the memory and the processor interconnected by a communication line as indicated in the diagram below:

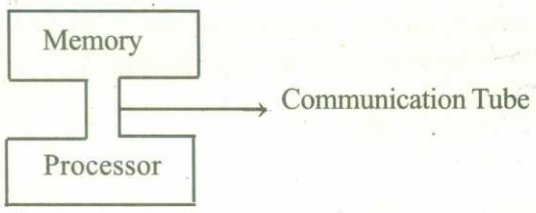


Fig. 1.1: Architecture of modern computer system

It must be mentioned here, that this concept of architecture assumes the use of Von Neumann model of computation and therefore cannot be generalised.

In 1977, Dizso Sima, *et al* interpreted the concept of architecture at a number of levels of increasing abstraction. At each level, the architecture has been defined by them stating the underlying computational model, the functional specification as well as the actual implementation. The interpretation covers three aspects:

- The underlying computational model;
- The level of consideration; and
- The scope of interest.

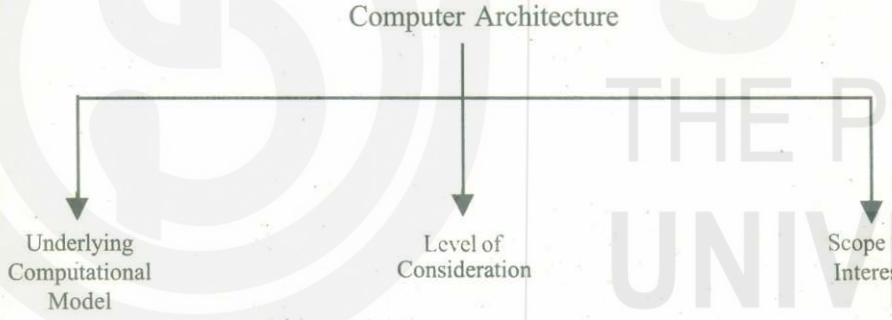


Fig. 1.2: Computer architecture

1.3.3 Basic Computer Architecture

The main components in a typical computer system are: the processor, memory, input / output devices, and the communication channels that connect them. The processor is the workhorse of the system; it is the component that executes a program by performing arithmetic and logical operations on data. It is the only component which creates new information by combining or modifying current information. In a typical system there will be only *one processor,* known as the central processing unit or CPU. Modern high performance systems often have more than one processor. Systems only with one processor are known as *serial processors,* or especially among computational scientists, they are known as *scalar processors.*

The second component is memory that simply stores information until it is requested by another part of the computer system. During normal operations it feeds instructions and data to the processor, and at other times it is the source or destination of data transferred by I/O devices. Information in memory is accessed by its *address.* In programming language terms, one can view memory as one-dimension *array* (M). A processor's request to the memory might be "send the instruction at location M (1000)" or a disk controllers' request might be "store the following block of data in location M (0) through M (225)".

I /O devices transfer information without altering it between the external world and one or more internal components. I/O devices can be secondary memories, for example, disks and tapes, or devices used to communicate directly with users, such as video displays, keyboards, and mouse.

The fourth component of the system namely the communication channels that tie the system together can either be simple links that connect two devices or more complex switches that interconnect several components and allow any two of them to communicate at a given point of time. When a switch is configured to allow two devices to exchange information, all other devices that rely on the switch are blocked, that is to say that they wait until the switch can be reconfigured.

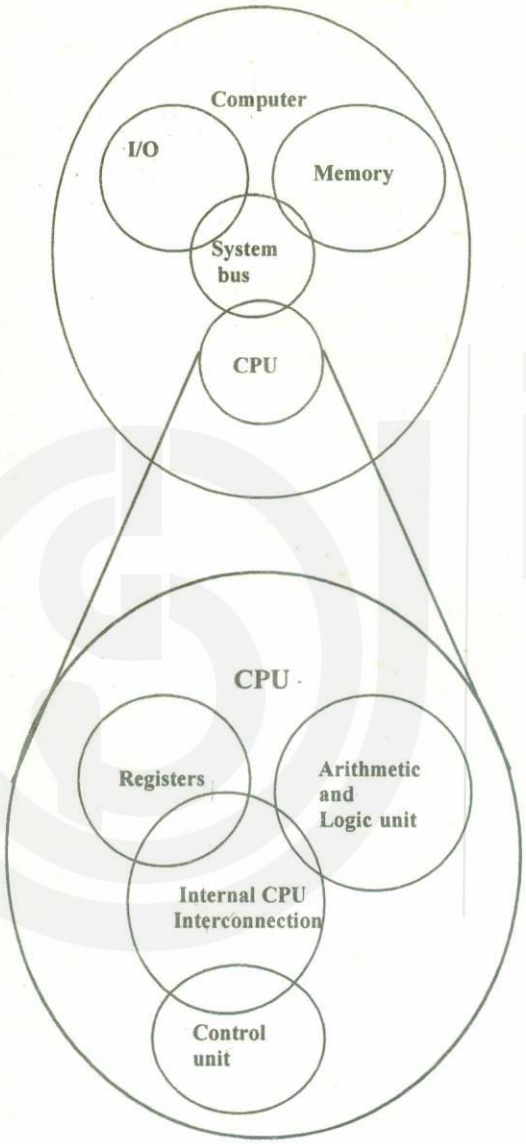


Fig. 1.3: The Central Processing Unit (CPU)

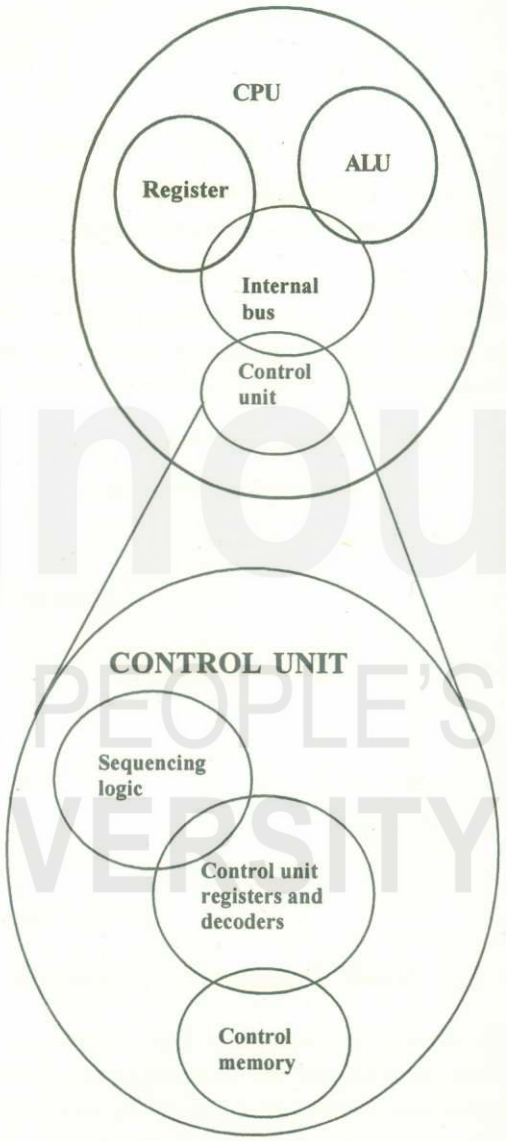


Fig. 1.4: The Control Unit

Source: Stallings, William (2003): Computer Organisation and Architecture, pp. 11-12.

There is a common convention used to draw simple 'stick figures' (diagrams) of computer systems known as PMS notation. In a PMS diagram, each major component is represented by a single teller, for example, P for processor, M for memory, or S for switch. A subscript on the letter distinguishes different types of components, such as Mp for primary memory and Mc for cache memory so on. Lines connecting two components represent *links* and lines connecting more than two components represent a *switch.* Although they are primitive and might appear at first glance to be too simple, PMS diagrams convey quite a bit of information and have several advantages, not the least of which is that they are independent of any particular manufacturer's notations.

3) What is meant by Processor Technology?

4) Explain briefly the two major categories of Microprocessor Technology.

5) Differentiate between the two concepts 'Computer Architecture' and 'Computer Organisation'.

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

1.4 STORAGE TECHNOLOGY

Storage technology is mainly concerned with computer memory system. Though, it may appear to be simple in concept, computer memory exhibits perhaps the widest range of type, technology, organisation, performance, and cost of any feature of a computer system. No one technology is optimal in satisfying the memory requirements for a computer system. As a consequence, the typical computer system is equipped with a hierarchy of memory subsystems, some internal to the system (directly accessible by the processor) and some external (accessible by the processor via an I/O module).

Two parameters characterise the storage technology: cost and speed of access which are directly proportional to each other. For a given capacity, the cost increases as the speed of access increases or the access time decreases. The storage in computer system is organised hierarchically with the cheaper and slower storage media at the top and the faster and expensive storage media at the bottom (see the diagram). The top-level storage is used for archival purpose and the bottom level to speed up computational process.

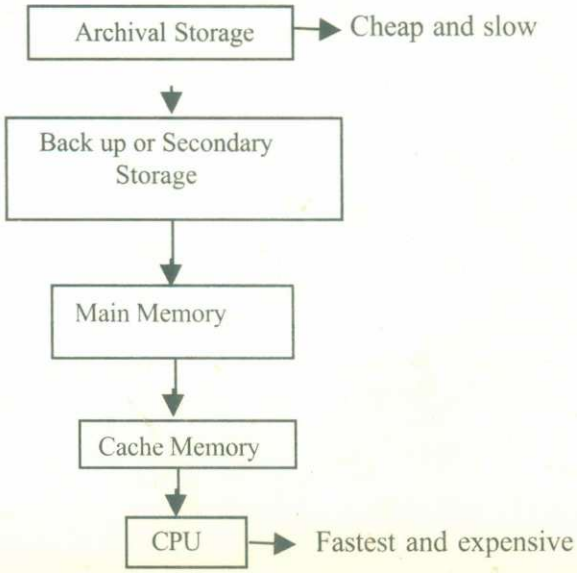


Fig. 1.5: Storage hierarchy in a computer system

The storage capacity increases as one moves up the hierarchy. The key characteristics of computer memory systems are indicated in the tablemm below:

Table 1.2: Key Characteristics of Computer Memory System

• **Location** **Processor** **Internal (main)** **External (secondary)** • **Capacity** **Word Size** **Number of Words** • **Unit of Transfer** **Word** **Block** • **Access Method** **Sequential** **Direct** **Random** **Associative**	• **Performance** **Access time** **Cycle time** **Transfer rate** • **Physical Type** **Semiconductor** **Magnetic** **Optical** **Magneto-optical** • **Physical characteristics** **Volatile / Non-volatile** **Erasable / Non-erasable** • **Organisation**

As might be expected, there is a trade-off among the three key characteristics of memory: namely *cost, capacity* and *access time.* At any given time, a variety of technologies are used to implement memory systems. Across this spectrum of technologies the following relationships hold:

- Faster access time, greater cost for bit
- Greater capacity, smaller cost for bit
- Greater capacity, slower access time.

The dilemma facing the designer is clear. He would like to use memory technologies, provide large capacity memory both because the capacity is needed and because the cost per bit is low. However, to meet the performance requirement, the designer needs to use expensive, relatively lower capacity memories with short access time. The way out of this dilemma is not to rely on a single memory component or technology, but to employ, a memory hierarchy. A typical hierarchy is illustrated in the figure.

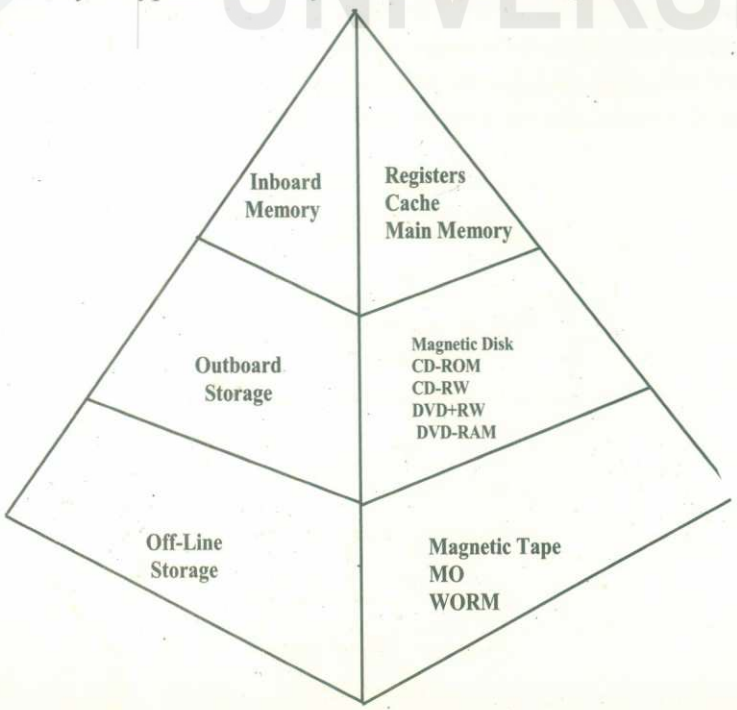


Fig. 1.6: The Memory hierarchy

Source: Stallings, William (2003): Computer Organisation and Architecture

1.4.1 Main Memory

In the early days, the main memory used to be built using magnetic cores. Currently cache and main memories are semiconductor memories. Cache uses the same technology as the CPU in order to a speed match between the two subsystems. Normally, semiconductor technology used is **transistor-transistor logic (TTL)** type for the cache memory. Main memory is made out of **metal oxide semiconductor (MOS) technology.** TTL is faster than MOS technology, consumes more power and is more expensive. Hence, large volume memories are not constructed using TTL technology.

There are two major categories of semiconductor memories: **read only memory (ROM)** and **random access memory (RAM).** ROM has a fixed content and is used for permanently resident programs and fixed data tables. RAM is used for transient program and data. Read only memories are of three types. First, ROM in which data is placed at the time of manufacture. The second type is programmable by the user. This type is known as **programmable ROM (PROM).** The third type is a **reprogrammable ROM** or **erasable PROM (EPROM).** The contents of EPROM's are erased by ultra violet light and reprogrammed electrically. It is not possible to erase EPROM locations selectively. The entire chip is exposed to UV light and all the contents are lost during such an exposure. Hence, reprogramming has to be done for the entire chip.

MOS is by far the most widely used technology for RAM. There are two major circuit divisions of RAM, viz. static and dynamic RAM (SRAM and DRAM respectively). In SRAM each memory cell contains transistors which are self locking into the 1or 0 states. But in DRAM each cell contains one or more transistors and a capacitor with the charge stored on the capacitor representing a 1 or 0. Because the charge on the capacitor leaks away with time it is necessary to refresh the DRAM periodically. SRAM has lower density on the chip than DRAM. Hence, higher capacity can be achieved in DRAM and hence they are cheaper than SRAM. Because of the refreshing requirement in the DRAM, it is slower than SRAM. The refreshing also calls for complex external circuitry. Both the SRAM and DRAM memory contents are destroyed if the power supply is switched off. Hence most computers have built in battery back up for at least a portion of the RAM.

1.4.2 Magnetic Storage

For over thirty years, magnetic storage has been essential part of the computer. Disks have been used as secondary storage medium and the tapes as archival storage medium. Disk memories have a capacity in the range of a few hundreds of kilobytes to a few gigabytes (one thousand megabytes constitutes one gigabyte, i.e. 10^9 bytes). One 'byte' consists of eight bits. Each of '1' and '0' is known as 'bit'. Disk memories fall under two broad categories: those with rigid recording media and floppies. With floppies the recording media can be removed and replaced whereas with the rigid media type, e.g. Winchester disk, the units are fixed and sealed. Rigid disks have access times of the order of 10 milliseconds for indexed data. For non-indexed data, the whole disk has to be searched which may take quite sometime, a few seconds to tens of seconds. Floppy disks revolve at slower pace and have access time of the order one or two seconds for the indexed data.

Magnetic tapes are typically used as archival storage. There is, however, increasing tendency to use floppies as archival storage for small volume data. For large volume data the present tendency is to use magnetic tape cartridges, which come with a capacity of 150 MB per cartridge.

1.4.3 Optical Storage

Optical storage technology offers an alternative to magnetic storage for both on line secondary storage and the offline archival storage. Optical storage technology eliminates some of the problems faced by magnetic storage technology. For example, to meet the

increased storage requirements of computers, the optical disks offer much higher storage density than magnetic disks. Optical storage does not suffer from the problem of long-term stability as the magnetic tapes do. Optical storage is like magnetic storage in many ways. It comes in disk, tape and card forms. Of these, disks are more popular. Disks come in many sizes: 12 inch, 8 inch, 5-1/4 inch and 3-1/2 inch. One of the computer products, CD-ROM (compact disk read only memory) has the same dimension as the CD in entertainment field from which it is derived. The performance of optical disks is broadly similar to that of magnetic disks. They serve the same application purposes as magnetic disks do. The fundamental difference being, data is retrieved using optical means in the optical storage technology. Data recording need not be by optical means.

Optical storage can be placed under three broad categories based on the data recording methodology:

1) Write once read many times (WORM)
2) Read only memory (ROM)
3) Rewritable memory.

The essential feature of WORM storage is that recording is irreversible; once the state of a storage element has been changed, it cannot be changed back to its original state. Data is written in sectors which are typically a few kilobytes in size. Sectors are treated as indivisible units because of the coding scheme used and hence a sector is written in full at a time. A sector can be superseded by a replacement sector located elsewhere. The data life is not infinite and the life reduces with the number of times it is read. This is the reason why the word 'many' is used in the description of the disk implying that the disk is not readable infinite number of times. Most of the manufacturers offer a shelf life of at least 10 years to 30 years or more for WORM disks.

The most popular of optical ROM in the computer field is the CD-ROM which stands for compact disk read only memory. In CD-ROM the content of the disk is a copy made in one operation for the entire disk, rather than copying sector by sector. Once copied, the contents cannot be changed in any way. There is no practical limit to the number of times the disk can be read.

The third class of optical disks are generally known as 'rewritable'. This is much closure to the magnetic counterpart in that data written on the disk can be erased and replaced by fresh data. Data is erased and written sector by sector; partial sectors cannot be changed. In the presently available system, two passes are required for replacing the information, first pass for erasing and the second for writing. Techniques which permit direct overwriting without the need for two passes are now being developed. One of the major hurdles in the development of the rewritable storage has been the 'fatigue' phenomenon exhibited by some media due to repeated erase and rewrite cycles. Research, however, continues to improve the number of erase/rewrite cycles that can be performed before the fatigue occurs.

Recording in the optical media could be analog or digital. Analog recording is essentially used for recording video or faster images like the ones that appear on the television. Digital recording is used for audio, character text, and bit mapped images. Digital recording may be done with or without error correction. For bit-mapped images, there may not be any need for error correction as there is considerable redundancy in original data itself. But for other textual matter, error detection/correction facility is required to ensure acceptable error rates from imperfect media.

Media management is an important aspect of all secondary and archival storage systems. Practically, all exchangeable magnetic media are loaded on to and removed from the drive by hand. The design efforts have been to simplify the loading/removal process for an untrained operator. Automated tape libraries using mechanical means specially designed

tape cartridges appeared in the 1970s but never widely accepted. One of the reasons for this is the non-compatibility of the tape cartridges used in automated systems with the tape spools used in manual systems. This led to data interchange and data portability problems. Hand loading is simple and widely used for optical storage too, but automatic library mechanisms seem to be more acceptable for optical storage than for their magnetic counterparts. These mechanisms referred to as 'jukeboxes'.

Optical disks used in jukeboxes are no different from those used in manually loaded drives and can be exchanged between the two. Usually, the disks are in the cartridge as used for manual loading, though some of the jukebox designs use bare disks as this take up less space to stack the multiple disks. Every jukebox has four components: a rack with slots for holding the disks, one slot for disk; one or more disk drives for data access; a mechanism for picking up a disk from the rack and load it on the drive and return the disk from the drive to the rack; and a controller which controls the jukebox operation based on the commands received from the CPU of the computer to which it is attached.

Self Check Exercises

6) What are the important characteristics of storage technology?

7) Explain the significance of storage technology.

8) Briefly explain three broad categories of optical storage technology.

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

1.5 INPUT / OUTPUT DEVICES

In order to use a computer to complete a given task, one must be in a position to instruct the computer by entering commands, instructions or information. This is accomplished through a variety of devices known as input/output devices. Input devices enable you to transform information into a digital representation that the computer can process. While on the other hand, output devices transform processed digital information into forms that humans can read and understand.

I/O devices can be broadly grouped as interactive and non-interactive devices. The most common example of an interactive device is the terminal with a keyboard and a display screen. Keyboard and mouse are examples of commonly used devices. Keyboard enables one to enter text into the word processing software so that one can complete his assignment, while the mouse enables us to navigate around the screen and select options by clicking the button. Once data has been entered into the computer, and it has completed its function you must be in a position to either see or obtain that information. This task is

accomplished by output devices. The most common output devices include the computer monitor and the printer.

In the following sections, let us try to learn about some of the important input output devices and understand their main characteristics.

1.5.1 Input Devices

This section attempt to explore the world of input devices, beginning with the most common ones:

Keyboards

Computer keyboards vary in design and layout but are all very much same in that they have 101, 104 or more *enhanced keyboards*. The main features included are Caps Lock, Num Lock, Function Keys, Ctrl, Backspace, Tab, Print Screen, and many other functions. The latest of keyboards have additional features such as specialised buttons that allow any person to navigate the Web, check e-mail, go to a home page, develop and use multimedia applications, search for files and facilitate the operation of one's music software with one touch. The most popular keyboard in use today is the *QWERTY Keyboard*. The operation of this keyboard can rarely attain speed over 100 Wpm (word per minute). Another type of keyboard that is used by many is called as the Dvorak keyboard named after its inventor. Many people claim that this keyboard is more efficient than QWERTY Keyboard. The main difference between these two keyboards is that in QWERTY system we have alphabets starting with Q then W then E and so on whereas in other system it starts with different alphabet.

Mouse

Mouse happens to be one of the most common input devices. There are different types of mouse. The commonly used mouse contains a rolling ball and several buttons that are pressed to execute commands. As the mouse is moved around a flat surface the ball rolls and feeds electrical signals back to the computer to move the cursor on the screen. One of the problems associated with this type of mouse is the dirt that is collected on rollers inside while it is used. The dirt will naturally effect the smooth movement of the mouse resulting in the halt of the cursor on the screen. Optical mouse is used to over come this type of problem.

Scanners

Print documents can easily be fed into the computer without retyping (keying) them manually. Scanners are the devices which accomplish this task. Scanners facilitate capturing of the information and storing them in graphic format for displaying back on the graphical screen. A scanner consists of two components, the first one to illuminate that to be scanned so that the optical image can be captured and the other component to convert the optical image into digital format for storage by the computer. Thus, the graphic images scanned can now be seen and processed directly by the computer.

There are different types of scanning devices but two of the most widely used ones are flat bed scanners and barcode scanners. Modern flatbed scanners are used to scan both text and images. When scanning the text/image, once it has been digitized, is then processed by the software so that each line of the image is converted into text. Pictures are scanned in a similar way. However, the scanner converts the image parts into digital form and divides the printed image into *a grid of small dots,* called *pixels*. Colours are scanned *pixel* by *pixel* and entered as graphic image into the computer. Advances in the technology have made scanners affordable and some of the computer systems provide a scanner along with the system.

The second variety of scanners is known as *bar code scanners*. These are small hand held devices that read the UPC (Universal Product Code) on merchandise. One of the

applications of these devices is to use in libraries to scan the accession number of books at the issue counters.

Graphic Tablets

Graphic Tablet is a device which is like an *electronic pantograph.* Pantographs were developed to trace drawings. The device consists of *two pens* connected by a mechanism. As one traces the drawing with one pen, the other pen moves and makes a copy of the original drawing. The *electronic graphic tablet* is simply a pad with an associated *pen-like stylus.* We move the stylus, across the pad to trace the existing drawing laid on top of the pad, or to create a new drawing. As we draw on the pad, the electronic signals are sent to the computer, the pad is divided into *pixels* and when we pass on a pixel we press on a pixel the corresponding pixel on the screen lights up.

Touch Screens

Some computers have touch screens that are sensitive to human touch. We use our finger tips to touch the screen and a command is executed. There are different devices used to make the computer screen sensitive to human touch, some of which are Infra red screens surround the screen with LEDs that emit infra red light. When one breaks the beam with finger the computer can detect it. There are also pressure sensitive screens that use two sheets of *mylar* assembled so that they are separated by a small space. When one touches the screen two mylar sheets come into contact thus completing an electrical circuit that is detected by the computer.

Voice/Speech Input

"It is stated that in spite of the incredibly fast processing speed of a computer, the keyboard slows it down". This is because to process information by computer you must first enter into the computer. People can only type with a definite speed, which is no match to the speed with which computer functions. The ideal situation, therefore, would be that people should be able to talk to a computer directly as they do with their colleagues. This would enable the computer faster and easily accessible to those who do not know how to type.

To enable humans to converse with the computer, Voice / Speech Input Devices have been developed. In fact, one of the most exciting areas of research is in recognising human voices / speech so that this could form input to computer directly. This approach will eliminate the need for keying in data. There are several problem areas for research since speech recognition system should be able to identify who is speaking and what the message is. Voice recognition techniques along with several other techniques to convert the *voice signals* to *appropriate words* and derive correct meaning of words are required for a commercially viable comprehensive speech recognition system. We have achieved some successes in this direction. Today devices are available commercially, to recognise and interpret human voices within limited scope of operation and application. Current systems can respond to a fairly small number of words or phrases from a person whose voice has been *'learned'*. Recognition rates are fairly high, but not yet adequate for a wide variety of applications.

Voice recognition is central theme in the development of fifth generation computers.

Digital Cameras

Digital photography is the term that applies to taking pictures without using film. The image is saved in a memory module on the camera and later on computer as a graphics file. The file then can be loaded directly into a graphics program where one can manipulate it and later print it. This approach to photography provides immediacy and ease of use which was not available previously in photography. This is achieved by the use of special devices known as digital cameras. Digital cameras function as input devices to computers.

To produce an image, a digital camera uses a small array of light sensors that creates an electrical charge when exposed to light. This is similar technology to that used in scanners, digital copiers and video cameras. The *number of picture elements (pixels)* in the sensor array determines the *optical resolution* at which pictures can be taken. For example, if a camera has optical resolution of 640x480 pixels, the array consists of 307,200 individual sensors that produce the final image. The total sensor array is called a *charge-coupled device (CCD),* and is used in most digital cameras. Digital cameras have made it easy to bring both images and video into the computer. Most of these devices are small and connect either through a USB port or a dedicated video connection. Some cameras allow to use a floppy disk and files are simply transferred via the floppy drive on to the computer. At present, the digital cameras are very costly. But as time progresses and technology advances, it is expected that the prices will come down and the digital cameras might become an attractive accessory with all computers.

Magnetic Ink Character Recognition

Magnetic Ink Character Recognition (MICR) is the process of reading characters which are printed in magnetic ink. The MICR devices are generally used by the banking industry to read the account numbers on cheques directly and do the necessary processing. These devices are relatively fast and error free but it can only accept numbers and 4 special characters. Cheques have a row of figures printed along the bottom in magnetic ink. This is done using an expensive and specialised machine which types the characters onto the document using a special ink containing *iron oxide.* The iron is magnetised and this magnetism can be read by the automatic equipment which sorts out the cheques and inputs the data from them.

Optical Character Readers

Optical Character Readers (OCR) make use of data input device which can recognise printed or typed characters by a light scanning process. Many commercial documents such as telephone bills, electricity bills, etc. have a row of figures across the bottom which are read by an OCR scanner. A special type face is used for the characters to enable them to be read easily.

A simple variation of OCR is mark sensing or Optical Mark Recognition (OMR). Mark sense equipment can recognise whether certain areas of a document have been shaded in pencil. The mark is sensed by using reflected light. Less light is reflected from the mark than from the back ground of the paper. Mark sense documents include UPSC examination application forms, examination answer sheets, foot ball pool coupons, forms used to record gas and electricity meter readings. These can also be used gainfully for marking multiple choice questions in examinations.

In addition to all these devices discussed in the foregoing paragraphs there are stand alone input devices like key-to-tape, key-to-disk and key-to-diskette. It may be stated that there are a range of input devices and one can select them as per the need. Let us now try to examine the range of output devices available and their main characteristics.

1.5.2 Output Devices

In the earlier sections we learnt about some of the major types of input devices and their important characteristics. In what follows, we shall attempt to know about output devices commonly used in the present day computer processing environments.

When you are working with a computer, you would like to see what you are working on. Images on a screen, are commonly referred to as soft copy because they are visible but they are still in digital form - which you cannot physically handle you might like to print your work on paper so that you can take it with you, or share it with others, or file it for

further use or reference. This kind of output is called a hard copy because it is physical and you have the ability to handle it. The devices that enable you to do this are called output devices. They are called so because they enable you to get information out of the computer.

The output normally can be produced in two ways; either on a display unit / device or on a paper. Other kinds of output such as speech output, mechanical output is also used in certain applications.

Display Devices

One of the most important peripherals in a computer is the graphic display device. Conventional computer display terminals known as alpha numeric terminals, display characters (images) from a multi-dot array (normally 5×7 or 7×9). These are used to read text information displayed on the screen. However, there are increasing demands for display of graphs, diagrams and pictures to make the visual presentation of information more effective for user interaction and decision making.

Graphic display is made up of a series of dots called '*pixels*' (picture elements) whose pattern produces the image. Each dot on the screen is defined as a separate unit which can be directly addressed. Since each dot can be controlled individually, there is much greater flexibility in drawing pictures. Three types of display screen technology has been identified and discussed in the literature of computer science. They are:

- Cathode Ray Tube (CRT)
- Liquid Crystal Display (LCD)
- Projection Displays

Cathode Ray Tube (CRT)

In CRT display terminal, the main component is the electron gun, the electron beam controlled by an electromagnetic field and a phosphor coated display screen. The electron gun emits an electron beam which is directed towards the phosphor coated display by the electromagnetic field in order to create an image. It may be mentioned that there are two types of CRT display - Vector and Raster Scan.

In vector CRT display the electron beam is directed only to the places where the image is to be created. On the other hand, in Raster Scan display the image is projected on to the screen by directing the *electron beam* across *each row of picture elements* from the top to the bottom of the screen. This type of display provides a high dynamic capability since image is continuously refreshed and it allows continual user input and output. It offers full colour display at a relatively low cost. Raster Scan displays are becoming more and more popular.

Liquid Crystal Displays (LCD)

These were first introduced in watches and clocks in 1970s. LCD is now applied to display terminals. The major advantage of LCD is the *low energy* consumption. The CRT is replaced by liquid crystal to produce the image. This does not have colour capability and the image quality is relatively poor. Liquid Crystal Displays are commonly used in portable devices which require low energy. Flat screens in computers are also in use these days.

Projected Displays

These are normally used for large group presentations. These systems can be connected to computer and whatever appears on the computer terminal *gets enlarged* and projected

on a large screen. These devices are commonly used in seminars, class rooms and marketing presentation.

Terminals

Terminals have become popular interactive input and output units. A terminal when connected to a CPU sends data and instructions directly into the computer. Terminals can be grouped into two types, namely *hard copy* terminals and *soft copy* terminals. A hard copy terminal provides a print-out on paper whereas a soft copy terminal provides a *visual display* on a screen. A soft copy terminal is also known as a CRT (Cathode Ray Tube) terminal. Terminals are also classified as *dumb terminals* or intelligent terminals depending upon provision for any intelligence or otherwise at the terminal. Current technology has brought about a low price differential between intelligent and dumb terminals, with the result; intelligent terminals are increasingly used in preference to dumb ones.

Graphic Display Terminal

A display terminal displays information in both character and graphic forms. These types of terminals are extensively used in CAD /CAM (Computer Aided Design and Computer Aided Manufacture) applications.

Printers

Printers are one of the most common forms of output devices. They are used for producing output on paper. There are a large variety of printing devices which can be classified according to the *print quality* and the printing speeds. Latest estimates indicate that there are about 1500 types of printers commercially available conforming to about 15 different printing technologies. Let us examine some of these.

a) Printing Technology — Impact Printers vs. Non-Impact Printers

Impact printers use variations of standard typewriter printing mechanism where a hammer strikes paper through inked ribbon. Non-impact printer uses *chemical, heat* or *electrical signals* to *itch* or *induce* symbols on paper.

b) Character form — Fully Formed Character vs. Dot Matrix

Fully formed characters are constructed from solid lines and curves like the characters of the typewriter whereas a *dot matrix* character is made up from a carefully arranged sequence of dots packed very close to each other. Obviously, print quality of a dot matrix character will be poorer compared to that from fully formed characters.

c) Printing Sequence — Serial vs. Line vs. Page

This indicates the amount of information a printer can output within *single cycle of operation.* SCM printing is done by character, whereas line printing forms an entire line and prints, and a page printer outputs a whole page of characters and images simultaneously *during one cycle.* From this it might be inferred that the speed of output will depend upon the *printing sequence* incorporated in the printing device. In other words, while character printers are slow in providing the output, page printers provide the fastest output and are also comparatively costly.

Let us now discuss two of the most popular printers.

Dot Matrix Printer

This is one of the most popular printers used for personal computing system. These printers are relatively cheaper compared to the other printers using other technologies. The Dot Matrix Printer (DMP) uses impact technology and a print head containing banks of wires moving with high speeds against inked ribbon and paper. Characters are produced in *matrix format.* The speed ranges from 40 characters per second (CPS) to about 1000 CPS. The disadvantage of using these printers is that the *print quality and speed* is low.

Laser Printers

The technology associated with laser printers is a high quality, high speed and high volume variety. It works in a *non-impact* fashion on plain paper. Printing is achieved by deflecting laser beam on to the photosensitive surface of a drum and the latent image attracts the *toner* to the image areas. The toner is then electro-statically transferred to the paper and fixed into a permanent image. Speeds can range from 10 pages per minute to about 200 pages per minute. This technology is relatively expensive but is becoming very popular because of the quality speed and *noiseless operations.*

Digital Plotters

Computer Aided Design (CAD) applications in areas such as *chip and PCB* design, engineering and architecture use digital plotters for output. These devices produce architectural plans, engineering drawings, chip layouts and maps. A digital potter has a pen, whose motion across the surface of the paper is controlled by a computer. Some of the models work with a range of colours. Digital plotters are very slow output devices and require special software to control them.

Speech Synthesis

Speech synthesis output is the counterpart of voice recognition input. Like voice recognition, speech synthesis is currently the subject of intensive research. A few speech synthesis output devices have been implemented, but the full potential of this technique is yet to be realised. Speech synthesis works by storing a digitally coded form of a *number of key sounds.* Words are constructed by combining these codes, and then *decoding the digital patterns* through a suitable set of circuits connected to a speaker. A different type of speech output is where key words and phrases have been recorded and the software causes these to be output in the appropriate order.

Computer Output on Microfilm (COM)

COM is an output technique gaining wide acceptance. COM is now used by banks for their daily records of account balances. It is also used in archiving records in museums, archaeological departments and libraries. It avoids the bulk an expense of the large quantities of paper produced by printers. A 'page' of output is displayed on a screen and photographed by a special camera. The film image of one page measures less than one fourth of an inch square. The film is cut into postcard-sized microfiches, which contain the images of approximately one hundred pages. A microfilm reader is used to project the enlarged image of a page on to a screen.

In the foregoing sections we have attempted to learn about a variety of input /output devices which are currently utilised in the activities relating to computerisation. However, it must be mentioned that the choice of these devices depends on a number of factors. Depending upon the circumstances, a suitable device that meets the requirements, has to be chosen.

Self Check Exercises

9) Identify the different categories of I/O devices for a computer.

10) Discuss the significance of speech synthesis output.

Note: i) Write your answers in the spaces given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

1.6 SUMMARY

An attempt has been made in this Unit to present an overview of the developments taking place in computer technology. At the same time it provides a window to know about the developments which have gone into generations of this technology. The areas discussed include computer generations, the field of processor technology, storage technology and input / output devices. The concept of computer architecture has been briefly discussed from a historical perspective bringing out the distinction between computer architecture and computer organisation.

It has been stated that the developments in the fields of processor technology, *specially microprocessors,* act towards reducing the 'speed gap' between the main memory and CPU. In storage technology the most recent advancement has been the development of optical storage. The optical storages have a variety, namely Read Only Memory (ROM) and Write Once and Read Many times (WORM) memory and Rewritable store. Unlike the magnetic counter parts, Optical storages are amenable for long time storage.

There are a variety of input / output devices available for the computer system. The most important of these is the interactive terminal with a keyboard and screen. Apart from this, modern devices like mouse, scanners, graphic tablets, light pens, touch screens voice input and digital cameras have been briefly discussed along with their important characteristics.

In describing the output devices care has been taken to highlight the significance of 'display devices', Liquid Crystal Displays (LCDs), graphic display terminals, laser printers, plotters, speech synthesis, and computer output on microfilm (COM), etc. The unit also includes features like self check exercises with model answers and a select list of references. These features are intended to further self learning process and to acquire additional information concerning the complex aspects of the subject discussed in the unit.

1.7 ANSWERS TO SELF CHECK EXERCISES

1) Technology is regarded as a specifiable way of doing any thing. In other words, technology is a procedure for solving problems. Information Technology (IT) is a recent term, which has become popular during the late 1970s. The term basically refers to technologies and applications developed through the combining data processing and storage power of computers with the distance transmission capabilities of telecommunications technologies. These two technologies were considered as

separate fields of study earlier. To be more precise, it is the convergence of computing and telecommunications based on microelectronics together with its various applications that people usually mean by Information Technology (IT). IT comprises five components: electronic and photonic technologies, computer technology, communications technology, artificial intelligence and human-machine interface technology. IT is used for information management, which involves: acquisition, processing, storage, retrieval, dissemination, display, reproduction and printing.

2) The chronological development of different generations of computers are:

1940s to early 1950s	1st generation Vacuum Tube based Bulky, consumed high power.
Early 1960s	2nd generation Based on transistor technology Smaller in size Consumed much less power More reliable than 1st generation.
Early 1970s	3rd generation Using Integrated circuits Reduced in the size Reduced power consumption Improved reliability Better human-machine Interface Multi programming facilities.
Early 1980s	4th generation Development of Small Scale Integration Medium Scale Integration Large Scale Integration Very Large Scale Integration.
Late 1980s	Emergence of Micro computers Super computer Parallel processing systems.
1990s	5th generation computers To handle sophisticated non-numeric applications such as Pattern-recognition, speech recognition, knowledge Processing etc. It includes Inference Engines, Knowledge- Bases, Natural Language Interface, Photonic devices and computing in optical domain.

3) Processor Technology: Processor is one of the main components of a computer system. It is a component that executes a program by performing arithmetic and logical operations on data. A processor consists of at least an instruction control unit and an arithmetic unit. In a typical system there will be only, one processor known as the central processing unit or CPU. Systems with only one processor are called *serial processors* or *scalar processors*. Modern computers have two part architecture consisting of memory technology and processor technology. The instruction register holds current instructions under execution and the program counter points to the next instruction. The important development is the introduction of cache memory between the main memory and Central Processing Unit (CPU). Cache memory concept is based on the assumption that operational access tend to cluster together for a short duration.

4) There are two categories of microprocessor technology: i) the traditional CISC (Complex Instruction Set Computing) technology and ii) the recent RISC (Reduced Instruction Set Computing) technology. The more instructions contained in Central Processing Unit (CPU) instructions set and the more complicated those instructions, the more *clock cycles* the CPU needs to complete an instruction search. Traditional CISC technology has complex sets of between 150 and 300 instructions and processes only *one instruction search* at a time.

On the other hand, the RISC chips contain perhaps 70 to 80 of only most frequently used instructions, and are capable of processing a new instruction on each clock cycle. RISC uses smart compiler technology, which not only translates commands from software language to machine language and sends them to the CPU, but does so, in the form of strings of the most frequently used instructions, broken down into

pieces for storing in a reserve memory cache. It also employs pipelining and multiple parallel processing to avoid internal bottle necks and to speed up processing.

5) In describing computers, a distinction is often made between 'computer architecture' and 'computer organisation'. Although it is difficult to give precise definition to these terms, a consensus exists about the general areas covered by each.

Computer architecture refers to those attributes of a system visible to a programmer or put another way, those attributes that have a direct impact on the logical execution of the program. In other words computer architecture is the study of the components that make up computer systems and how they are interconnected. On the other hand, computer organisation is concerned with the implementation of computer architecture. It refers to the operational units and their interconnections that realise the architectural specifications. As an example, it is an architectural design issue *whether a computer will have a multiply instruction.* It is an organisational issue *whether that instruction will be implemented* by a 'special multiply unit' or by a 'mechanism that makes *repeated use of add unit* of the system. The organisational decision may be based on *the anticipated frequency* of use of the 'multiply' instruction, the relative speed of the two approaches, and the cost and 'physical size of the special multiply unit'.

6) Computer memory is also known as storage. The technology associated with memories is also referred to as storage technology. Memory cost and speed of access which are directly proportional to each other. In other words, for a given capacity, the cost increases as the speed of access increases or 'the access time' decreases. Stated differently, the performance of a memory system is defined by two different measures 'the access time' and 'the cycle time'. Access time, also known as 'response time' or 'latency', refers to how quickly the memory can respond to a 'read' or 'write' request. Several factors contribute to access time of a memory system. The main factor is the physical organisation of the memory chips used in the system. Memory cycle time refers to the minimum period between two successive requests. These are some characteristics relating to storage technology.

7) Though seemingly simple, the concept of computer memory exhibits the widest range, type technology, organisation, performance and cost of any feature of a computer system. No one technology is optimal in satisfying the memory requirements for a computer system. As a consequence, the typical computer system is equipped with a hierarchy of memory subsystems, some internal to the system (directly accessible by the processor) and some external (accessible by the processor via an I/O module).

Computer memory is organised into a hierarchy. At the highest level (closest to the processor) are the processor registers. Next comes one or more levels of 'cache'. When multiple levels are used, they are denoted by L_1 and L_2 etc. Next comes the main memory, which is usually made out of dynamic random- access memory (DRAM). All these are considered internal to the computer system. The hierarchy continues with external memory, the next level typically being a fixed hard disk, and one or more levels below that consisting of removable media such as optical disk and tape. As one goes down the memory hierarchy, one finds decreasing cost / bit , increasing capacity, and slower access time. It would be nice to use only the fastest memory, because that is most expensive memory, we trade off access time for cost by using more of the slower memory.

8) Optical storage can be broadly grouped into three categories:

- Write once read many times (WORM)
- Read only memory (ROM)
- Rewritable memory

The essential feature of WORM is that recording is irreversible; that is once the state of storage element is changed, it cannot be changed back to its original state. Data is written in sectors which are typically a few kilobytes in size. The data life is not infinite and the *life reduces* with the *number of times it is read.* Generally the shelf life is at least 10 years and in some cases it may be 30 years or more for WORM disks.

The most popular implementation of ROM in the computer field is the CD-ROM which stands for compact disk read only memory. In the contents of the disk is the copy made in one operation for the entire disk rather than copying sector by sector. Once copied contents cannot be changed in any way.

The third category of optical disks is generally known as 'rewritable', because data written on the disk can be erased and replaced by fresh data. Data is erased and written sector by sector. Partial sectors can be changed. The process of rewriting is a two stage process the first pass for erasing and the second for writing. The recording can be either analog or digital.

9) **Input Devices**

- Keyboards
- Mice
- Scanning devices
- Graphic Tablets
- Light Pens
- Touch Screens
- Voice Input
- Digital Cameras

Output Devices

- Display Screens
- Graphic Display Terminals
- Cathode Ray Tubes
- Liquid Crystal Displays (LCDs)
- Printers
- Plotters
- Speech Synthesis
- COM

10) One of the most exciting areas of research in computer science is 'teaching computers to understand spoken communication'. As the voice input continues to be viewed as a future technology considerable progress is being made in this direction.

Speech synthesis output is the counter part of *voice recognition input.* Like voice recognition, speech synthesis is currently the subject of intensive research.

A few synthesis output devices have been implemented, but the full potential of this technique is far from being realised. Speech synthesis works by storing digitally coded form of a number of key sounds. Words are constructed by combining these codes, and then decoding the patterns through a suitable set of circuits connected to a speaker. Another type of speech output is where *keywords and phrases have been recorded* and *the software causes* these to be output *in appropriate order.* Voice recognition techniques along with several other techniques to convert the voice signals to appropriate words and derive the corresponding meaning of words are required for this purpose. We have found limited success in this area and today devices are available commercially to recognise and interpret human voices within limited scope of operation. Obviously, this method eliminates the need for keying in data.

1.8 KEYWORDS

Arithmetic and Logic Unit (ALU) : A part of the computer that performs arithmetic operations, Logic operations, and related operations.

Assembly Language : A computer oriented language whose instructions are usually in one-to-one correspondence with computer instructions and that may provide machine language output in terms of '0' and '1'.

Associative Memory : A memory whose storage locations are *identified by their contents,* or by a part of their contents, rather than by their names or positions.

Cache Memory : A special buffer storage, smaller and faster than main storage, that is used to hold a copy of instructions and data in main storage that are likely to be needed next by the processor and that have been obtained automatically from main storage.

CD ROM : Compact Disk Read Only Memory. A non-erasable disk used for storing computer data. The standard system uses 12-cm disks and can hold more than 500 Mbytes.

Instruction Set : A complete set of the operators of the instructions of a computer together with the description of the types of meanings that can be attributed to their operands. Synonymous with machine instruction.

Input Devices : A device by which data can be entered into a computer system.

Output Devices : A device used for transferring data from internal storage of a computer to an external device or user.

Input-Output (I/O) : Pertaining to either input or output, or both. Refers to the movement of data between a computer and a directly attached peripheral.

Information Technology : The acquisition, processing, storage and dissemination of vocal, pictorial, textual and numerical information by a micro-electronics based combination of computing and telecommunication.

Storage Device : In computing, a unit into which data can be entered, retained and later retrieved.

1.9 REFERENCES AND FURTHER READING

Amdahl, G.M., *et al* (1964). Architecture of IBM system/360. *IBM Journal of Research and Development*, 8(2), 87-101.

Bradley, A.C. (1989). *Optical storage for computers: technology and applications*. Chichester: Ellis Horwood Limited.

Ceruzzi, P.E. (1998). *A history of modern computing*. Cambridge, M.A.: MIT Press.

Clemens, A. (2000). The undergraduate curriculum in computer architecture. *IEE Micro*, May /June 2000.

Halal, William E. (1993). The information technology revolution: computer hardware, software and services into the 21st century. *Technological Forecasting and Social Change*, 44, 69-86.

Hennessy, J.L. and Patterson, D.A. (1990). *Computer architecture: a quantitative approach*. N.P: Morgan-Kauffman.

Hussain, Khateeb M. and Hussain, Donna (1989). *Computers: technology applications, and social implications*. New Delhi: Prentice-Hall of India.

Hwang, K. (1993). *Advanced computer architecture*. New York: Mc Graw-Hill.

Martin, James (1989). *Information engineering: introduction*. New Jersey: Prentice-Hall.

Martin, William J. (1995). *Global information society*. 2nd rev.ed. London: Aslib. pp. 33-63.

Saffady, William. (1992). *Optical storage technology 1992: a state of the art review*. Westport: Meckler.

Sima, Dezso, *et al* (2003). *Advanced computer architectures: a design space approach*. Delhi: Pearson Education Inc.

Smith, R.E. (1989). A historical overview of computer architecture. *Annals of the History of Computing*, 10, 277-303.

Stallings, William. (2003). *Computer organisation and architecture: designing for performance*. 6th ed. Delhi: Pearson Education Inc.

Vranesic, Z. and Thurber, K. (1980). Teaching computer structures. *Computer*, June 1980.

ignou THE PEOPLE'S UNIVERSITY

UNIT 2 SOFTWARE

Structure

2.0 Objectives

2.1 Introduction

2.2 Software

2.3 Functions of Software

2.4 Types of Software

2.4.1 System Software

2.4.2 Application Software

2.5 Markup Languages

2.6 Summary

2.7 Answers to Self Check Exercises

2.8 Keywords

2.9 References and Further Reading

2.0 OBJECTIVES

In the previous Unit you have learnt about hardware aspects of Computer Technology. In this Unit we will be discussing the software aspects which are the key component of Computer Technology.

After studying this Unit, you will be able to:

- know what the concept 'software' means;
- interpret the basic functions of software;
- list different concepts related to software as per their hierarchical relationship;
- understand different system software—operating system mainly;
- gather ideas regarding differences between a single and Multi user OS;
- develop your knowledge about different application software;
- clarify the important points related to Bibliographical DBMS; and
- above all to understand fundamentals of software.

2.1 INTRODUCTION

Since its inception the computer has two inter-dependent concepts – hardware and software. To tell about a computer we broadly mean electronic circuitry within a cabinet, screen, and a keyboard. These are called hardware, which has physical existence and one can touch it as an object, can feel its presence. The term 'hardware' denotes the machinery or equipment components of the computer system. They are visible and tangible. But a computer cannot do anything on its own; it is deaf and dumb unless supported by some instructions. More specifically a computer requires sequentially arranged instructions to do a desired job. Such a sequence of instructions written in a language is called a 'program'. It is the program that controls the activities of processing by the computer and the computer performs the job. These exist within purview of software. The term 'software' denotes the programs or predefined sequences of instructions that a computer

follows while executing a task. Mostly people do not feel the existence of software as it does not have any physical shape as an object. Similar analogy can be seen in case of electrical wiring at our home, we get light from a bulb without feeling the continuous flow of electricity over the wire. However, this unit is to give an overview of the basics of computer software.

2.2 SOFTWARE

Software, in general, is a set of programs that direct computers to perform desired functions. In simple terms, a software program is a set of instructions for a computer. In other words, it refers to all programs that are written to be executed on a computer. The programs are written in number of different languages, called programming languages. Conventional software programs are maintained on auxiliary storage media for transfer to the CPU's RAM whenever required. Documentations to understand program's logic and execution are also integral part of the software.

The software is an intellectual product and is neither physically visible nor tangible. The software component of the computer is the predetermined program instructions which are maintained on a magnetic or optical storage medium from which it is loaded into the CPU's RAM for execution. In some computer systems these software packages are loaded by the manufacturer as a part of the specifications of the computer units.

2.3 FUNCTIONS OF SOFTWARE

The main function of software is to make use of hardware to perform a desired job. For carrying out different jobs, different types of software are designed and developed. They are mentioned category wise with their specific functions in the next section. For example, general purpose application software are for jobs like word processing, spreadsheets, graphics, etc., whereas specific purpose application software are for specific tasks like payroll, income tax, library management, etc. Similarly System software are responsible for handling System control programs which include operating systems, database management system, communication monitors; System support programs which include utilities, job accounting, etc.; and System development programs which include compilers, interpreters, assemblers, etc.

Self Check Exercises

1) What is software?

2) What are the functions of software?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

2.4 TYPES OF SOFTWARE

There are differences in opinion in grouping software. Some are in favour of making four basic types of it. According to them, these are System software, Application software, Utility software and Software Package. System software is a collection of programs that manages all the concurrent tasks being performed by a computer, including the execution of application software programs. Application software performs specialised functions like library automation and the MIS functions mentioned in an organisation, or other useful work not related directly to the operation of the computer itself. Utility software is a set of programs that perform routine day-to-day tasks, such as listing or compressing data, copying files, etc. Software Package is computer application consisting of a set of programs created to perform a particular type of work. Software packages are generally grouped into Modules. Each module is supposed to handle a particular task independently. There is integration at some stage to have interlinking and information flow. For example, in library management system the software package may contain modules like Catalogue management, Circulation, Journal Subscription, Membership, New Arrivals, Budget Monitoring, etc. Many people want to keep mainly two types System and Application Software.

2.4.1 System Software

System software is a set of one or more programs, designed to control the operation of a computer. System software makes the operation of the computer system more effective and efficient. In general, system software supports smooth running of the other software; communication with peripheral devices; supporting the development of other types of software; and monitoring the use of various hardware resources. In brief, it provides environment to develop and operate application software. There are different types of system softwares: a) Firmware, b) Operating System, (c) Translator, (d) Simulators; and (e) Emulators.

Firmware

When a computer is started, the CPU requires a few instructions and data stored in ROM of the microprocessor. Such instructions are essential as it includes the way of initialising other peripheral chips, way of establishing communication with the keyboard, etc. Frequently used program instructions which the computer executes are permanently stored in ROM chips. These programs are available to CPU immediately on permanent basis whenever it requires. These programs are categorized as 'firmware'. Without the firmware no computer can work as the boot up programs that are written to start the computer are also in the form of Firmware.

Operating System

An operating system acts as a master control program on the computer. It controls the flow of signals from CPU to various components. In other words, operating system is a set of programs which control the computer and enable it to perform various tasks, including communicating, via commands, with the user. It permits the computer to supervise its own operations automatically calling in the applications programs and managing the data needed to produce the output desired by users. Operating systems help in performing memory management, file management, program interaction, communication networking, graphics management and other utilities. Operating systems can be divided into the following types: Batch Processing System, Multitasking System, Time Sharing, and Multiprocessing System.

a) ***Batch Processing Systems***

This operating system accepts jobs (user requests) and executes them on a serial basis. The jobs are placed in a queue before execution. In this case, the data are collected over several days or weeks in a transaction file and then processed all at one time, as a batch against a master file. Periodic processing of payroll system is an example of such systems.

b) ***Multitasking System***

This operating system can perform two or more task concurrently on the same computer with one CPU. For example the system may accept commands for writing a report with one program while another program may search a database or bring out a print copy of a processed item. The best application of it can be seen nowadays in microcomputers on window-based environment. Although the CPU processes one program at a time, it gives a feeling that it processes all the programs simultaneously at a time. It is due to the processing speed which is very fast. The human reaction time is generally 0.1 second and CPU responds to user in less than this time unless there is some execution problem.

c) ***Multiprogramming System***

The term Multiprogramming System is used mainly for Multi-user environment, but when designed for personal computers, they are called concurrent processing operating systems. Multiprogramming is the concurrent execution of two or more programs assigned by different user terminals on a multi-user operating system. Like multitasking systems, here also the CPU spends a certain amount of time for executing each program one after another.

d) ***Time Sharing System***

A time-sharing system accomplishes direct access to a computer's CPU by a number of independent users (or terminals) simultaneously. Several users can use the resources of a single computer on a time dependent framework. The time slices which may range from 10 to 20 milliseconds are allocated to each. The difference between multitasking and time sharing systems lies in the fact that in multitasking systems the CPU directs the programs to take turns by providing a different amount of time as required to accomplish them whereas in time sharing systems the CPU spends a fixed amount of time (or time slice) with each program before going to the next one.

f) ***Multiprocessing System***

Multiprocessing involves simultaneous processing of two or more programs by more than one computer at a time. This is in contrast to multiprogramming in which two or more programs are processed concurrently by a single computer. In multiprocessing the computers are linked together to carry out the multiple programs.

Self Check Exercises

3) What is system software?

4) What are different types of Operating System?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

Single User Operating System : These Operating Systems allow input device single communication with the CPU at a time; hence the name is Single User Operating System. MS-DOS is well known example of this type.

MS-DOS (Microsoft Disk Operating System): MS-DOS is a single user operating system which was developed by Microsoft, USA in 1981 for IBM-compatible PCs. IBM version of this operating system is also known as PC-DOS. MS-DOS is more popular for its text oriented user interface. User has to type commands at the command line using the keyboard to get the computer's response

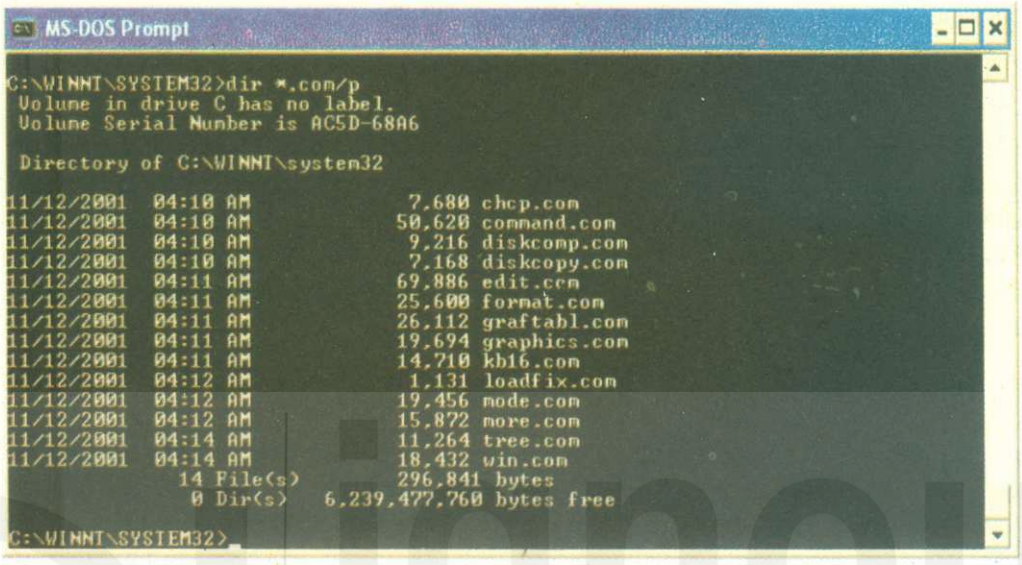


Fig. 2.1: MS-DOS Screen

The commands are grouped into two: internal and external commands. Internal commands (viz. DIR, DEL, COPY , etc.) are within a single separate file COMMAND.COM, called command processor and are loaded into the RAM at the time of booting. Hence they are executed directly from RAM. External commands (viz. CHKDSK, EDIT, FORMAT , etc.) are the names of the programs stored as an executable disk file. So they must be loaded into the RAM when needed. With the release of a new version, it is a usual fact that some commands are withdrawn sometime and some new commands are introduced. Above figure shows result of DIR command execution.

WINDOWS : Windows is an operating environment developed by Microsoft Corporation for a large variety of IBM-compatible machines with Intel microprocessors. Instead of typing commands as in DOS, it provides a Graphical User Interface (GUI) which presents programs, procedures, files, commands, etc. as graphical symbols which the user can select by pointing to with a mouse and then clicking one of its buttons. When Windows is used with DOS, the user gets the facility of graphical interface instead of typing the commands. A window is a rectangle that appears on the screen and displays information for some specified purpose. The operating system (or the operating environment) which is window based, can display several windows simultaneously on the screen each meant for different purpose and may belong to different programs like word processing, spreadsheets, etc.. However, conventionally Windows (capital 'W') is an operating environment developed by Microsoft that lays a graphical user interface sheet around the MS-DOS or PC-DOS operating system. Windows makes use of windows extensively for its functioning.

Today there are many versions of Windows in widespread use. Windows 2000, Windows XP, Windows 2003 versions are popular these days. Each version is an improvement over the previous version though the basic features of Windows remain same. Several windows can be displayed on the same screen. Windows also facilitates more than one program to run at a time and each program to share information with other programs.

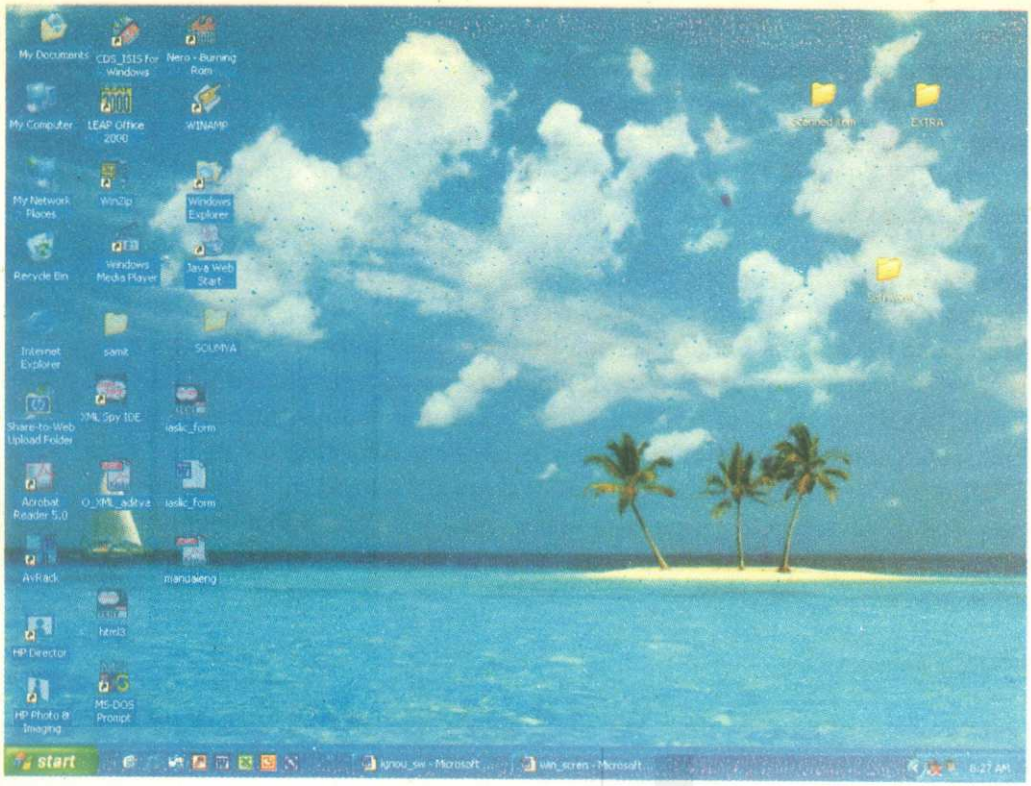


Fig. 2.2: Windows Screen

Start button and taskbar are located on the bottom of the screen when windows start for the first time. By default they are always visible when Windows is running. Taskbar provides a log of all programs that are opened. To switch programs, the buttons on the taskbar indicating the programs are to be clicked. Above figure shows display of a typical windows startup screen.

Self Check Exercises

5) What are different types of commands in DOS?

6) What is Windows?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

Multi User Operating System: These Operating Systems allow input devices more than one communication with the CPU at a time; hence the name is Multi User Operating System. UNIX, LINUX are well known example of this type.

UNIX

UNIX is a multitasking and multi-user operating system developed at the Bell Telephone Research Laboratories, a division of AT & T Information System in 1969. It was developed

for larger machines. As the major part of the operating system was written in C, a high level language, it became the first open operating system that could be loaded on to any platform from PCs to Super computers. Unfortunately, this advantage of UNIX led to many variations in UNIX. There are about 18 major flavors of UNIX — SCO-UNIX, Solaris, UNIX-OSF/1, LINUX, etc.. The Microsoft version of this operating system is also known as XENIX.

UNIX is a popular operating system because it can be rapidly implemented on any computer system. There are two basic versions of UNIX-one is called System V from AT& T and the other is Berkley version from Berkeley University. The flavor of UNIX used by Silicon Graphics on the Indigo computers is actually a mixture of both the System V and Berkeley versions of UNIX. UNIX is a line-oriented system similar to MS-DOS. Each command to the operating system is terminated by the 'return' or 'enter' key. UNIX is case sensitive and uses the ASCII character set and supports upper and lower case. The disadvantage of using both upper and lower case is that commands and file names must be typed in the correct case. Most of the UNIX commands and file names are in lower case.

UNIX Layers

UNIX consists of three layers — the UNIX Kernel, the UNIX Shell and UNIX utilities/ Application layer. These layers are developed from protection point of view and each layer is meant for specific functions.

The UNIX Kernel: This is the most important and the innermost layer. It has a process scheduler, a hierarchical file structure and mechanisms for processes to communicate with each other.

The UNIX Shell: This is the middle layer which contains command interpreter. It executes users commands and programs.

UNIX Utilities/Applications layer: This is the outermost layer containing programming tools and application programs. The programming tools include assemblers, compilers, editors, etc. The application programs include packages for database management, word-processing, graphics, etc.

Linux

Linux (pronounced as Lih-nuks) is a free operating system that was created by Linus Torvalds when he was a student at the University of Helsinki in 1991. Linux is neither a public domain, nor a shareware. It is *free* software, commonly called *freeware* or *Open Source Software*. One can give away or sell copies, but inclusion of the source code is mandatory. It follows POSIX (Portable Operating System Interface for UNIX) standards. So Linux looks like UNIX, but does not come from the same source code.

Linux is only the kernel of the operating system to control hardware, to manage files, to separate processes, and to do many similar jobs. There are several combinations of Linux with sets of utilities and applications to form a complete operating system. Each of these combinations is called a *distribution* of Linux. Like other operating system, Linux is made up of software instructions that lie between the computer hardware (disks, memory, ports, and so on) and the application programs (word processors, Web browsers, spreadsheets, and so on). At the center is the kernel, which provides the most basic computing functions (managing system memory, sharing the processor, opening and closing devices, and so on). Besides the kernel, Linux provides the following other basic services:

File systems: The file system provides the structure in which information is stored on the computer. Information is stored in files, primarily on hard disks inside the computer. Files are organised within a hierarchy of directories. The Linux file system holds the data files, the programs to run, and the configuration files to set up the system.

Device drivers: These provide the interfaces to each of the hardware devices connected to the computer. A device driver enables a program to write to a device without needing to know details about how each piece of hardware is implemented. The program opens a device, sends and receives data, and closes a device.

User interfaces: An operating system needs to provide a way for users to run programs and access the file system. Linux has both graphical and text-based user interfaces. GNOME (GNU Network Object Modeling Environment) and KDE (K Desktop Environment) provide graphical user interfaces, whereas shell command interpreters (such as bash) run programs by typing commands and options.

System services: An operating system provides system services, many of which can be started automatically when the computer boots. In Linux, system services can include processes that mount file systems, start network, and run scheduled tasks.

As an alternative to Windows, the Linux version of UNIX is growing in popularity. At least one office suite (Corel) is now available for the Linux platform. The advantage of Linux is that it is widely available, well supported, and inexpensive (or free). It is also less demanding on hardware resources. Linux uses Internet and industry standard components and protocol giving a system with complete network integration. The operating system can act as a server for most major file serving protocols and provide all the major Internet applications. Linux is normally obtained as a 'distribution'. This is a combination of the Linux operating system kernel and other tools, utilities and applications. Some of these are available for free over the Internet and others on CDROM. Because Linux itself is free software that can be freely copied, many distributions are available both over the Internet and sold on CD-ROM with added convenience and support.

Network Operating System : The Network Operating System is the main system software to work in a network environment. And most of the recent computers have the facility for network connections. A network is a group of computers connected to each other or to a central server so that they can share resources such as documents and printers. A network can have a variety of platforms having different systems. In addition, a server can act as a client to another server. A machine designated as a client, however, cannot be a server. There are many operating systems that can be used as NOS.

It also helps to gain access to the Internet and send/receive messages by using e-mail or help in connecting office and home computers. This is accomplished by the Network Neighborhood icon which appears on the desktop. By double-clicking this icon, one can get an overview of all the computers within the workgroup or any Netware servers to which the computer is connected.

In UNIX based LAN environment, the TCP/IP is used as the network protocol. Each system on the network is uniquely assigned a four-byte IP address, just as in the case of Internet, which uses mostly TCP/IP. The clients, if they are not UNIX machines, use TCP/IP clients. One such program for DOS based PCs is PCTCP/IP, a freeware from National Centre for Supercomputing Application (NCSA). However, to use X-windows based programs, one has to have an X-emulator on PCs like Xvision or Deskview. As network under activity is with TCP/IP, UNIX can be extended to be on Internet. In fact, UNIX is the most popular OS on Internet, where the servers can be used as LAN and Internet servers. In fact, LINUX can also be used in Network environment.

Windows NT

Windows NT from Microsoft is gaining popularity basically for two reasons. It provides windows environment and it is cheaper. Of late, it is gaining popularity over Novell Netware, although a feeble contended to UNIX.

Developed by Microsoft in 1993, Windows NT (NT stands for Network Technology) is an operating system which provides multi-user, multitasking features with graphical user

interface. It is mainly intended to support large network of computers. Unlike the early Windows operating environment (which ran with DOS), Windows NT is a true operating system, interacting directly with the hardware. Windows NT is compatible to most modern PCs with powerful microprocessors of both Intel's Pentium and Motorola's PowerPC. It provides client/server-working mode, which is used in network computing system, in which the client is the requesting machine and the server is the supplying machine. The server usually offers services to large number of computers on the network known as clients. Windows NT is easy to install and use. Windows versions (such as Windows 2000, ME, XP, etc.) developed after Windows NT are based on Windows NT technology.

Netware

This product is from Novell Incorporation. Originally Netware was designed to support a network of DOS machines. A recent version of Netware allows a wide variety of Operating Systems on the LAN, and is extended to be an Internet server. The latest version Netware 5 supports Novell's own IPX protocol and TCP/IP. Its strong competitor is Windows NT.

Self Check Exercises

7) What do you understand by the layers of UNIX?

8) Write down the common features of Linux?

9) What is an open source?

10) What do you mean by NOS?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Machine Level Language Generators

i) **Translator:** Translator is meant to translate one language to another. So a translator is mainly related to computer language. In case of a computer, its hardware part only can operate when instructions are made of 0s and 1s, i.e., in machine language. But it is not that easy for human being to remember them correctly. Thus a move towards substituting such instructions of machine code by letter symbol-mnemonics was taken. Mnemonic codes ease programmer to write efficient program. But an intermediatory agent is required to translate these mnemonic codes into machine

codes. There are different types of translators for different categories of languages viz. assembler for assembly language, interpreter and compiler for high-level language.

ii) ***Assembler:*** When an assembly language program is required to run or execute on a computer, the program is needed to be converted into its equivalent machine language program. The mnemonic operation codes are to be translated to absolute numeric operation codes. The translator which does this conversion is known as assembler. As the assembly language is machine dependent, assembler is also a machine dependent system program which must be supplied by the computer manufacturer. Hence, an assembler program translates source program in assembly language into object program in machine language.

iii) ***Interpreter:*** High-level languages are machine-independent to facilitate execution of a program on different computers. Machine-independence is achieved by using different translators to translate a high-level language program to machine languages of different computers adopting one-to-many translation techniques. All high-level language translators consist of three distinct phases viz. lexical analysis, syntactical analysis and semantic analysis. There are two types of translators for high-level language program – interpreter and compiler. Interpreter is one which translates one statement of source program at a time into one or more machine instructions (i.e., object code) subject to prompting programmer to debug, if any, the source program and then instructs the computer to perform immediately those instructions before the control moves towards the next statement. In other words, an interpreter is a program which translates one statement of a high-level language program into machine code. An interpreter is a smaller program as it is not required to create a permanent object module as in the case of a compiler. Hence it requires less memory space to work. The main disadvantage of interpreter is that every time one needs to execute the program, it is to be started from the very beginning. Thus it requires more time to execute on a computer.

iv) ***Compiler:*** Compiler is one which is used to store the high level language program, scan it and translate the whole program at a time into an equivalent machine language program that is a compiler translates the source program producing a complete machine language program/object program. In other words, a compiler compiles a high-level language program such as FORTRAN, C++, COBOL, etc. to convert it into a machine language program. The compilation of high-level language program is a complex process and consists of broadly two phases – analysis phase and synthesis phase. In the analysis phase, the source program is analyzed for its lexical, syntactical and semantic structure. The second phase is the synthesis of the object program in the machine language. The compiler checks all kinds of limits, ranges, errors, etc. Once the result of checking is satisfactory and error free, it produces a complete machine language program of minimized length ready for execution. This executable program, stored as a separate file, can be transferred alone with a storage media like floppy, CDROM from one machine to another where it can be executed without the presence of the compiler at all. A compiler takes more time to compile than to interpret, but a compiled program runs much faster than an interpreted program.

v) **Device Driver:** Devices are peripherals attached to a computer for different purposes. For examples, a printer is attached for producing hardcopy output, so it is known as output device, similarly, CDROM, Scanner, etc. A specialized program is required to establish a communication among hardware of such device, hardware of computer and the respective application software so that the device works properly. They act more or less as interface programs. These programs are called device driver. There are two more relevant concepts – Simulators and Emulators. A stimulator is a device, data processing system, or a computer program that represents certain features of the behaviour of the system. Emulators are the hardware or software designed to cause one device to behave as if it were another.

Self Check Exercises

11) What do you mean by Interpreter?

12) What is a compiler?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

2.4.2 Application Software

Application software performs specialized functions or other useful work not related directly to the operation of the computer itself. Application Software Packages are computer application consisting of one or more programs created to perform a particular type of work. It is designed to satisfy the needs of many organisations. Examples are Adobe PhotoShop, MS-Word, Lotus 123, Excel, Page Maker, CorelDraw, PowerPoint, MS-Photo Editor, FoxPro, dBase, CAD, etc. Some people have again liked to get a separate category for a set of programs that perform routine day-to-day tasks, such as listing or compressing data, saving file from virus infection, etc. This may be called as utility software or system support program/software. The utility software is so designed that it can perform maintenance work on the system or its components. Therefore, it is also known as Service Software. Win Zip, PC Tools, Norton Utility, etc are examples. Different types of application software are discussed in the following sections.

Data Base Management System

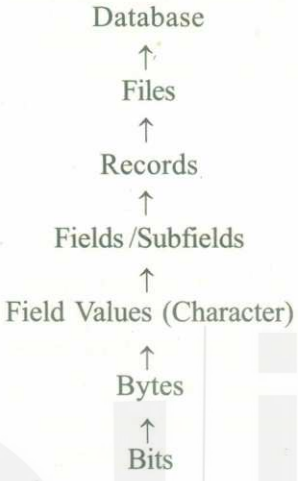
a) ***The Database Concept***

Data may be defined as a set of characters or symbols, which represent facts about certain objects, ideas or values. By its own, data represents a set of characters only that does not make any sense. On processing, meaning is attached to data, which transforms into information. For example, the digits 09271972 does not convey any meaning upon presentation, these are just a group of symbols, which may represent some idea. But again these are not perceivable unless one knows how to process it. If it is told that mmddyyyy format is followed here, i.e., the first two digits (mm = 09) represent the month, the next two (dd = 27) represent a month and the last two (yyyy=1972) represent a year, then it is easily understandable that this particular group of symbols stand for the 27th day of September in the year of 1972.

Thus the signs and symbols, which make up any data, represent certain ideas. But that idea may be conveyed only if the symbols are processed in the proper way. The idea, contained in any data, if conveyed explicitly, is termed as information. Hence, the entire information industry depends on the proper management of data, and management of data means their proper storage retrieval, up gradation, maintenance and processing. For ease of management, individual but related data-items are grouped together – thus forming databases.

b) ***Database***

Databases are organised collection of logically related but physically separated data files consist of data arranged in a structured form, designed to meet the information requirements of the user. With the help of a database questions can be answered, decisions can be made resources are identified and located. Librarians and Information Specialists have always been database experts. But the databases they deal with, generally give references to other sources of information or documents and are called bibliographic databases: Databases can be grouped into numeric and nonnumeric or textual database; nonnumeric database again can be divided into full text and bibliographic database.



In the computer world we frequently deal with files, which are the outer identifying boundary or a sort of folder containing data. Thus a database is collection of physically separated but logically related files; a file is a collection of similar records; a record is a collection fields/sub fields.

c) ***DBMS***

A database management system (DBMS) is a suite of programs which typically manage large structured sets of persistent data, offering adhoc query facilities. A DBMS can be an extremely complex set of software programs that controls the organisation, storage and retrieval of data (fields, records and files) in a database. It also controls the security and integrity of the database. The DBMS accepts requests for data from the application program and instructs the operating system to transfer the appropriate data.

When a DBMS is used, information systems can be changed much more easily as the organisation's information requirements change. New categories of data can be added to the database without disruption to the existing system.

Data security prevents unauthorised users from viewing or updating the database. Using passwords, users are allowed access to the entire database or subsets of the database, called subschema (pronounced 'sub-skeema'). For example, an employee database can contain all the data about an individual employee, but one group of users may be authorised to view only payroll data, while others are allowed access to only work history and medical data.

The DBMS can maintain the integrity of the database by not allowing more than one user to update the same record at the same time. The DBMS can keep duplicate records out of the database; for example, no two customers with the same customer numbers (key fields) can be entered into the customer database. The DBMS has been discussed in details in the Course 6, Unit 3.

Word Processor

A word processing package (or program) is the software which is used to format, create, edit, print, store and retrieve text material. The text material may be letters, reports, notes, theses, books, invoices, projects or anything else. A word processor is the software that performs the word processing task. Word processors facilitate faster production of documents as compared to traditional typewriters. Although the textual matter is typed into the computer through the keyboard, which resembles the traditional typewriter, the keyboard has certain special (functional) keys which the typewriter does not have. The task of editing the textual matter; adding and deleting of sentences, paragraphs; setting of margins; spelling checking, etc. are much easier and almost simultaneous in word processing rather than typewriting. Therefore, long documents, repetitive letters, memos, and general correspondence are the most widely used applications for word processors. The most significant advantage of word processors is increased productivity.

Desk Top Publishing (DTP)

Desk Top Publishing is now a full-fledged technology. It has become one of the fastest growing sectors of personal computing applications. Often Word Processing (WP) systems are used to assist DTP by creating ASCII text file, prior to typesetting. There is a very little distinguishing measure between DTP and WP. However, DTP can be defined as little more than the logical extension of word processing and as time passes WP packages has gradually adopted more and more DTP features. Slowly these two different application software packages are merged and become a single document production package.

It has the following new concepts:

a) Seeing and Painting for easy placement of text;

b) Progressive Disclosure for providing only relevant information;

c) Uniformity of commands across domains to make system procedures consistent; and

d) WYSIWYG (What You See Is What You Get) to make display a close facsimile of the final product.

To look back at the history and development of DTP it is understood that in 1980's Paul Brainerd, President, Aldus Corporation had developed a powerful word processor called page maker. Apple computer in order to push their mackintosh combined with Aldus and offered a package consisting of a Mackintosh, Laser Writer and Page Maker which Brainerd called a 'Desk Top Publisher (DTP)'. Since then many manufacturers including IBM have joined the bandwagon to propagate the concept. In 1987, DTP came to India when an Apple computer, Mackintosh along with Laser writer and Page Maker was first displayed at the PAMEX- 87 printer's exhibition. Later, the name was changed as Desk Top Publishing.

Components of DTP System

It has already been mentioned that DTP system is built with some necessary components having specific objectives. These components of DTP system are: Input Devices, Page Processing, and Output Devices. Input to DTP system can come from many sources. Any word processor which produces a standard ASCII text file can be used to develop the text component of DTP. Graphics can come into the DTP system - that is, digitized either through previously digitized clip-art, through one's own digitizing of photographs or line-art using any paint software. Finally, Scanners, which are non-keyboard input devices, can scan typewritten pages and enter the typescript into the computer as a text file. This

is done using OCR (Optical Character Recognition) package. Now-a-days a good number of software is well equipped with the page processing technique which involves use of Page Composition Software and Page Description. For output devices cost is a good guide to quality. For example, the Laser Jet Plus costs less than the Laser Printer or Laser Writer Plus. Decreased dpi (dots per inch) means lower resolution of graphics and typographic elements.

Advantages of DTP

The reasons which probably are the driving force for moving in the direction of DTP are:

a) ***Low cost Publication***

Laser printers which produce pages with high resolution (that is, more dpi) than before have come within the affordable price range of many. With increased dpi, letters can be smoothed and graphics can be produced at a higher resolution. Thus, the possibility of having more attractive publications is high. Besides this, due to continuing decrease of costs in computer memory and micro-CPUs, one can digitally manipulate typographic features that were not possible at all in the past. For example, the Postscript Page Description Software allows letter rotation that was not possible before from even the most expensive mechanical typesetting equipment.

b) ***Minimum Time Requirement***

Much of the typesetter's time has traditionally been spent composing input material from typewriters into pages with fonts, leading columns, etc. Now authors can do a great deal of that composing work with the word processor. Hence publishing time goes down.

c) ***Flexibility in Qualities***

This is the heart of the personal publishing revolution. With a desktop system, there is no need to imagine how the publication will look and suffer disappointments when the final result does not measure up to expectations. A text designer can choose Page Description Software Languages such as Postscript to produce a document with 600 - 1200 dpi on the Laser printer. Further, in each case, typefaces and fonts can be changed instantly. Thus, the level of publication quality solely depends on the designer and the budget.

d) ***Convenient for Designers***

Digital manipulation of words and paragraphs is probably the most important convenience provided. The case of correcting errors allows a more interactive writer/revise process. With WYSIWYG capabilities, one enjoys in page design, layout and typography by reviewing the entire piece as it is composed, rather than isolated reviewing of single words and sentences.

e) ***System Integration in Office***

Library applications publishing system can be integrated with PC networks and E-mail system. With such networks applications images received via E-mail can be incorporated in page.

Self Check Exercises

13) What is Word Processor? What do you mean by formatting documents?

14) Write down the advantages of DTP.

Note: i) Write your answers in the space given below.

ii) Check your answer with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

Presentation Software

Presentation software is a special type of graphics (or graphically presented text) which is used to exploit effects of movie. MS-PowerPoint is a presentation tool which helps to create eye-catching slides. The slides can be displayed electronically (using a computer), or in standard 35mm format (using a slide projector), or printed on overhead transparencies or paper. It also helps to bring out printed handouts for presenting a lecture and create notes to be used as references for effective and authentic presentation.

a) ***PowerPoint views***

Microsoft PowerPoint has three main views: normal view, slide sorter view, and slide show view. Any one of these main views can be used as default view in PowerPoint.

i) Normal views

Normal view is the main editing view, which is used to write and design the presentation. The view has three working areas: on the left, tabs that alternate between an outline of slide text (**Outline** tab) and slides displayed as thumbnails (**Slides** tab); on the right, the slide pane, which displays a large view of the current slide; and on the bottom, the notes pane.

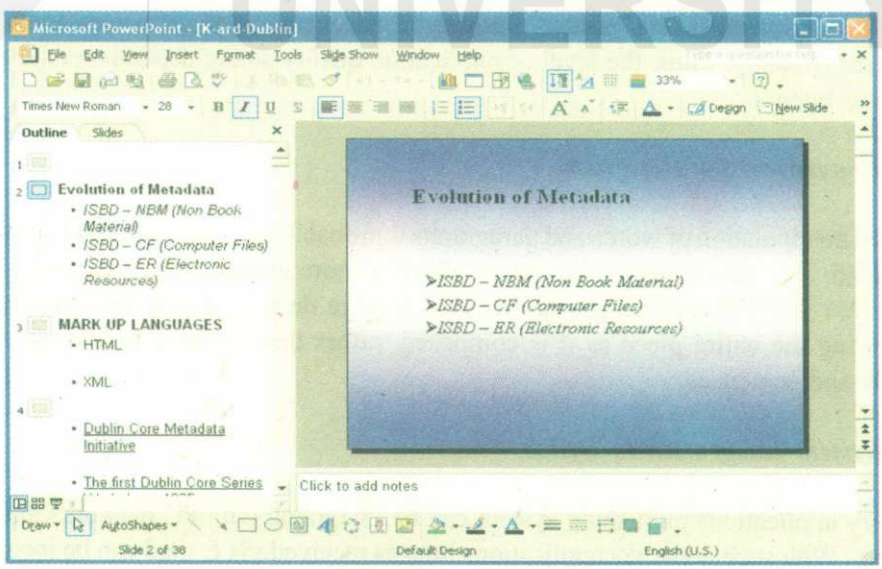


Fig. 2.3: Power Point (Normal View)

The *Outline tab* shows the slide text in outline form, this area is a great place to start writing the content — to capture ideas, plan of presenting them, and moving slides and text around. The *Slides tab* is to see the slides as thumbnail-sized images. The thumbnails make it easy to navigate through the presentation and to see the effects. The *Slide pane* shows the current slide in a large view, one can add text, insert pictures, tables, charts,

drawing objects, text boxes, movies, sounds, hyperlinks, and animations. The *Notes pane* is to add notes that relate to each slide's content. The size of the different panes in normal view can be readjusted by dragging the pane borders.

ii) **Slide Sorter View**

Slide sorter view is an exclusive view of your slides in thumbnail form. When you are finished creating and editing your presentation, slide sorter gives an overall picture of it — making it easy to reorder, add, or delete slides and preview your transition and animation effects.

Fig. 2.4: A View of Power Point Slide Sorter

iii) **Slide Show View**

Slide show view takes up the full computer screen, like an actual slide show presentation. In this full-screen view, a complete presentation is displayed with graphics, timings, movies, animated elements, and transition effects will look in the actual show.

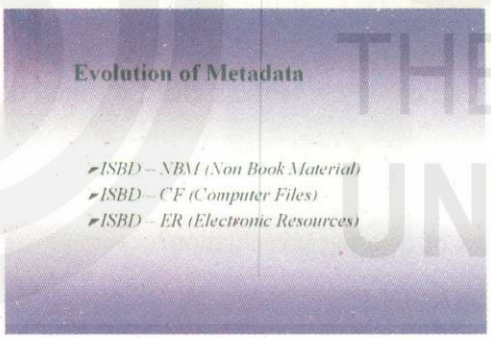


Fig. 2.5: A View of Power Point Slide Show

b) ***Creating Presentations***

Creating a presentation in Microsoft PowerPoint involves starting with a basic design; adding new slides and content; choosing layouts; modifying slide design, if you want, by changing the colour scheme or applying different design templates; and creating effects such as animated slide transitions. The *New Presentation* task pane in PowerPoint gives a range of ways with which to start creating a presentation. These include:

i) **Blank Presentation:** In this case, slide starts with minimal design and no colour. But presentations with colour scheme, fonts, and other design features with a wide range of predefined slide formats and Clip Art graphics libraries can be implemented. These predefined slide formats help to create slides quickly based on standard layouts and attributes applied to them.

ii) **Existing Presentation**: New presentation is prepared based on already written and designed one. This command creates a copy of an existing presentation so the design or content may be changed as required for the new presentation.

iii) **Design Template:** Presentation is made based on a PowerPoint template that already has a design concept, fonts, and colour scheme. Some of the templates come with PowerPoint, and some can be borrowed from the web site also. The Auto-Content Wizard is used to apply a design template that includes suggestions for text on slides. Microsoft.com helps to select templates from the Template Gallery which is arranged according to type of presentation.

iv) **Other sources:** Slides content from the other sources of presentation or text (e.g. MS Word) can also be inserted. OLE facilities are also available.

c) ***OLE (Object Linking and Embedding)***

The PowerPoint allows the OLE facilities to the presentation. The main differences between linked objects and embedded objects are where the data is stored and how the data is updated after it is placed in the destination file. When an object is linked, information is updated only if you modify the source file. Linked data is stored in the source file. The destination file stores only the location of the source file and displays a representation of the linked data. When an object is embedded, information in the destination file doesn't change if you modify the source file. Embedded objects become part of the destination file and, once inserted, are no longer part of the source file.

d) ***Modifying Visual Elements***

Changing the colour scheme : After creating a presentation, its background or colour can be changed to meet the specific requirements. The colour scheme dialog box can be used to select the colour for one slide or a series of slides at a time. Of course, colours other than the ones provided in the colour scheme can also be used if required.

Modifying the Slide Master : Every design includes a Slide Master, which identifies the position of the text and graphic objects, the style of the fonts, the footer elements, and the background of all slides. Any modification is brought to the Slide Master; it will be reflected in each slide that is based on the Master.

Adding objects: After selecting a particular slide layout, PowerPoint permits to add any object, i.e., text, clip art, a table, a chart, a sound or video clip or any other object to it.

Self Check Exercises

15) What are the ways of creating presentation?

16) What do you mean by OLE?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

An electronic spreadsheet is a computer generated worksheet which consists of a number of rows and columns to be filled with any set of data. Electronic spreadsheets (or simply spreadsheets) are useful tools for large tabulated computations and data analysis. They are used in several fields which deal in numbers or quantitative figures. The main functions of spreadsheets are:

- To help in tabular presentation of any quantitative data for computation,
- To help in recalculation of data easily reflecting changes, if any, in the existing data set,
- To perform simple and statistical calculations,
- To help in determining relationships, if any, among data sets,
- To help in creating budgets, financial reports, forecast cash needs,
- To facilitate graphical presentations.

Lotus 123 is a well-known electronic spreadsheet though it is an integration of three-spreadsheet, DBMS and graphics. Now-a-days, MS Excel, an important component of MS Office package, is popularly used electronic spreadsheet.

Features

The display screen of an electronic spreadsheet is divided into rows which are numbered numerically and columns which are labeled alphabetically. In case of Excel, a worksheet has 256 columns and 65,536 rows. The intersection of a column and a row is called a cell. The cell is named by combining the letter of the column and the number of the row, for example C6 means the intersection of the column C and row 6. Such names are called the cell address.

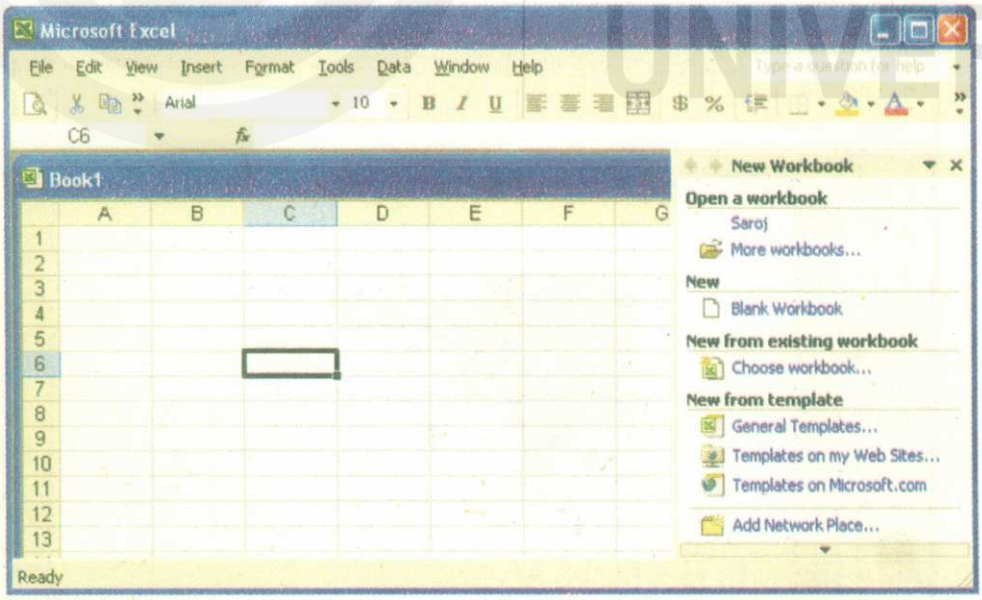


Fig. 2.6: A view of Spread Sheet

A cell can contain either text or number. To enter a continuous text adjacent cells can be used together. The cell on which data is entered is called active cell. Three types of data are entered into the cells: *Label*: A set of characters that are used for descriptions, headings and titles. Normally, a label begins with an alphabet. *Value*: Numerical data those are to be processed or analyzed. *Formula*: A relationship among cell values or a function of those.

Self Check Exercise

17) What are the functions of an Electronic Spread sheet?

Note: i) Write your answer in the space given below.

ii) Check your answer with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

Graphic

Adobe Photoshop software introduces facilities for image editing with powerful new features that offer something for every user. With built-in vector drawing tools and enhanced layer controls, Photoshop greatly expands your creative options. For Web graphics, no other program offers the comprehensive range of tools as is provided by Photoshop and its advanced Web-processing component, Adobe Image Ready™ 3.0 software. Interface enhancements help beginners learn more quickly and experienced users broaden their skills. Used around the world to create the highest quality images for Web sites, brochures, presentations, and much more, Adobe Photoshop is the centre of the digital-imaging universe. With its comprehensive set of retouching, painting, drawing, and Web tools, Photoshop helps to complete any image-editing task efficiently. And with features like the History palette and editable layer effects, you can experiment freely without sacrificing efficiency.

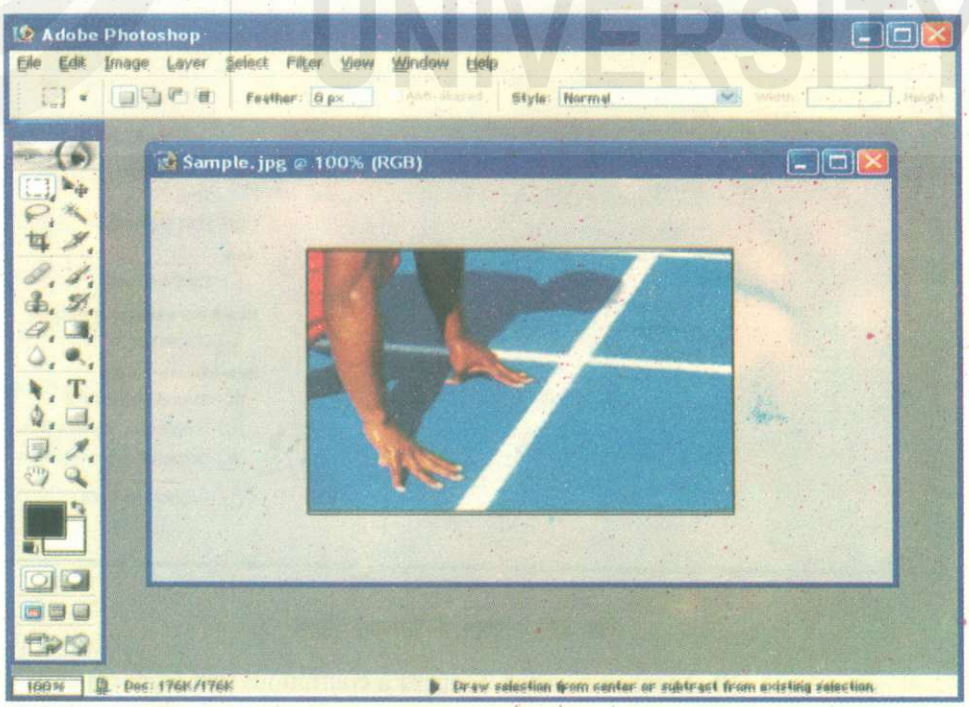


Fig. 2.7: A view of graphics

Photoshop offers high-powered image editing, photo retouching, and compositing tools to help in getting the best professional-quality results. CorelDraw and some other graphic software also offer similar kinds of facilities.

An animation is a sequence of images, or *frames*, which is displayed over time. Each frame varies slightly from the preceding frame, creating the illusion of movement when the frames are viewed in quick succession.

An image can have multiple animations that are associated with different rollover states. For example, adding an animation to the Normal rollover state causes the animation to play when the Web page is first loaded by a Web browser. Adding an animation to another rollover state causes the animation to play only when the Web user performs the specified action (such as placing the mouse over the rollover or clicking on the rollover).

Animations that you view in a Web browser are called *animated GIFs*. When you save an optimized document containing an animation, you can choose to generate an HTML file that contains code for displaying the animated GIF in a Web page. The resulting Web page can contain just the animated GIF or additional Web features, such as links and rollovers, depending on the source document.

You can apply optimization settings to animated images just as you do to non-animated images. You should always optimize an animation in GIF format because GIF is the only format in ImageReady that supports the display of animated images on the Web. While you can optimize an image that includes animation in JPEG or PNG format, these formats do not support animation. The resulting Web page will display only the current frame of the animation.

In addition to the standard optimisation options for GIF format, you can optimize frames to include only areas that change from frame to frame. (This greatly reduces the file size of the animated GIF.) ImageReady also applies a special dithering technique to animations to ensure that dither patterns are consistent across all frames and to prevent flickering during playback. Due to these additional optimisation functions, ImageReady may require more time to optimize an animated GIF than to optimize a standard GIF.

Audio/Video Software

Audio/Video is now available in digitized form and played with a computer. Songs or speeches we listen from tape recorder or radio are in analogue form. Digitization of audio can be done by attaching an audio player to a system through an audio capture card and then sound can be recorded to the system using various formats like .wav, .mp3, .midi, etc. Video capture also requires recording through video capture card with input from video cassette player or movie camera. The digitized file can be saved in .mov or .avi , divx, .mpeg.

Microsoft Windows Media Player is easily available with Microsoft Windows. It is a universal media player one can use to receive audio, video, and mixed-media files in most popular formats. Use Windows Media Player to listen to or view live news updates or broadcasts of your favourite sports team, to review a music video on a Web site, to 'attend' a concert or seminar, or to preview clips from a new movie. Windows Media Player is found to be the most comprehensive, simple, and efficient multimedia player. They provide simplified playing of a wide range of file types, high-quality multimedia experience, quick access to media content and can be easily tailored to meet your preferences.

Multimedia Authoring Software

Integration of audio, video, graphics, animation and text is known as multimedia. Software devised for production and presentation of complex, inter-related media objects has got extra importance and is called multimedia-authoring tools. Hence, a multimedia-authoring tool is the means that allows multimedia presentation, which synthesizes content created by the text, graphics, video and audio editing tools into a single presentation. There are broadly four categories of such software.

Card or page based tools : Multimedia elements are organized in this case as a page of a book or stack of cards. This is best used when bulk of content consists of elements viewed individually where hyperlinks enable users to browse in a non-linear way. For example, *HyperCard, SuperCard, MediaObjects, ToolBook*, etc.

Icon based tools: In this case, multimedia elements and user interaction are organised as objects in a structural framework. These tools actually help in organising the content by displaying a flow diagram of activities along with branching paths. The best examples are, *AuthorWare, IconAuthor, Quest*, etc.

Time based tools: Events and elements are organised along a timeline. These tools are best used when message or information must be provided within fixed time limit. An example of this kind is, *Macromedia Director*.

Object oriented tools: Here elements and events become objects that live in a hierarchical order — parent and child relationships. This is particularly useful for games, which contain many components with many personalities. The best examples are *mTROPOLIS, Apple Media Tools, Media Forge*, etc.

The eight mostly common features are important for any multimedia authoring software: Editing feature, Organising feature, Programming feature, Interactive feature, Performance tuning feature, Playback feature, Delivery feature, Cross platform feature.

The authoring environment is one of the most important areas of consideration. Basically authoring environments come in two varieties — (1) icon-based with drag-and-drop flowcharting, (2) object-oriented and require scripting.

The authoring tools must be able to handle rich text, graphics, animation, video, and sound. Hypertext and equation editing capabilities may also be necessary. The development tool must be able to handle different file formats depending on the media being used in the application. The following table presents the media types and the various file formats the content may be stored. For different media types, the following are the format.

Image	.BMP, .DIB, .PCX, .TGA, .GIF, .JPG, .TIF, .PICT
Sound	.WAV, .MID, .MP3, .SND
Animation	.FLI, .FLC, .MMM
Video	.MPG, .MOV, .TGA, .AVI

Authorware Attain

Authorware Attain is a leading visual rich-media-authoring tool for creating multimedia based Web pages and online learning applications. It allows developers, instructional designers, and subject matter experts to develop trackable learning applications and deploy them across the Web, LANs, and CD-ROM.

Director

Macromedia's Director is a powerful authoring tool for multimedia. Director enables users to import, integrate, and manipulate media objects to create high-impact interactive multimedia titles that can be delivered on Windows, Macintosh, and the Web.

Flash

Macromedia's Flash is useful in designing and delivering low-bandwidth animations, presentations, and Web sites. It offers scripting capabilities and server-side connectivity for creating attractive applications, Web interfaces, and training modules. It is very easy to create small & fast Shockwave multimedia, create animations or integrate existing graphics into your Web site, with no download lag.

Besides these three a good number of authoring packages are available in the market. Though they have many common and some unique features. For example, Authorware is good for multimedia based computer aided teaching. Macromedia Flash is good for special web application. Authorware and Flash require less technical skill whereas Director requires higher technical skill.

Paint is a drawing and painting software supplied with Windows. One can use to create simple or elaborate drawings. These drawings can be either black-and-white or colour, and can be saved as bitmap files. These drawings can be used for desktop background, or it can be pasted into another document. Paint is also used to view and edit scanned photos. One can use it to work with pictures, such as .jpg, .gif, or .bmp files. You can paste a Paint picture into another document you've created, or use it as your desktop background.

The following major tasks are frequently performed with Paint :

- To draw horizontal, vertical, or diagonal lines
- To use foreground and background colours to fill areas of graphics/drawings
- To use a picture created in Paint as the background for desktop
- To make detailed changes to the design and colour of an image by displaying gridlines.

With a digital camera or a scanner connected to computer, one can acquire an image from that device, work on the image in Paint and after editing or working on an image, it can be sent to friends, business associates, or anyone else using e-mail.

Communication Software

With the advent of networking, email and above all Internet, the communication software came into existence. Many software of such category are available. Internet explorer is one of them.

a) Internet Explorer

Internet Explorer is developed at the National Center for Supercomputing Applications based on NCSA's Mosaic technology. This communication software makes it easier to handle the World Wide Web, whether searching for new information or browsing a favourite Web site. Built-in technology saves time in completing routine Web tasks by automatically completing Web addresses, detecting network connection status, etc. A few basic things about browsing the Web are essential to understand for browsing the Web easier and faster. To start browsing the Web, click the Internet Explorer icon of the desktop and type the Internet address in the Address bar, then either press enter key or click the *Go* button; this will display the desired web page on the screen. However, one can use a favourite page or home page (e.g. www.ignou.ac.in) as default in the Internet Explorer properties and the page appears Internet Explorer is invoked.

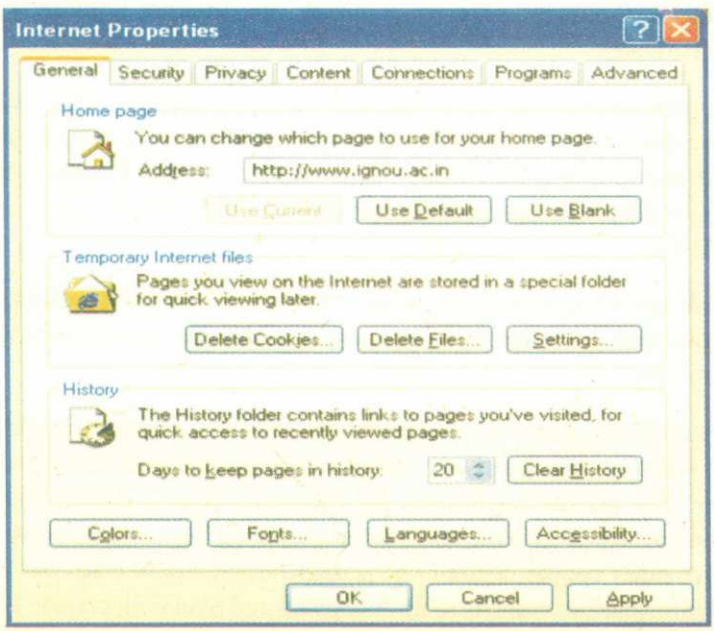


Fig. 2.8: Internet Properties

One can check whether any item on a page has hyper-link by moving the mouse pointer over the item. If the pointer changes to a hand, the item has a link to other page. A link is generally represented either by a picture, by a 3-D image, or by coloured text (usually underlined).

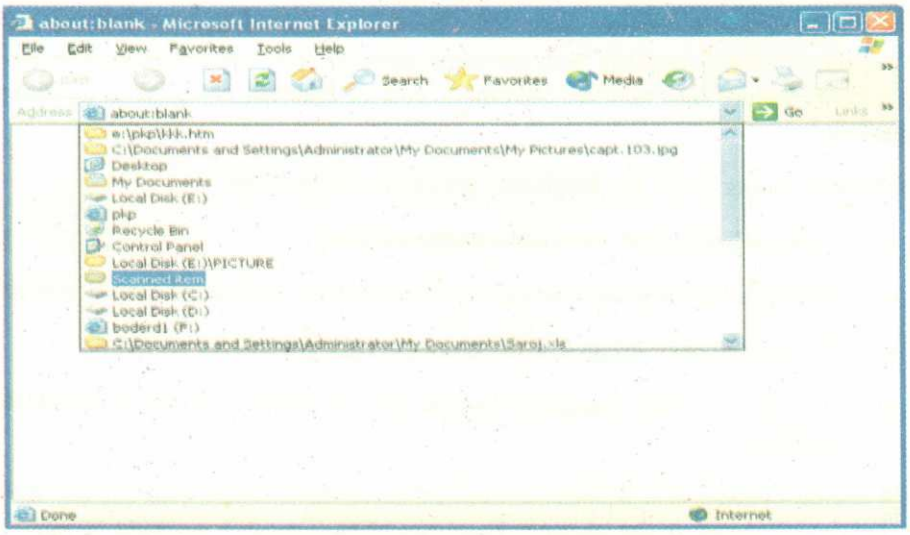


Fig. 2.9: Internet Explorer

However, at the address bar an URL (Uniform Resource Locator) is used. URL can be used for an Internet address to open a Webpage (e.g.www.iaslic.org), a program name with path to run that program (e.g. C:\MSOffice\Winword\Winword.exe) or a folder path to browse that folder (e.g. C:\Recycle Bin). Many self-explanatory options are available with the tool bar, for example, Back, Forward, Refresh, Home, Favourite, Search, History, etc.

Security and Privacy

In Internet where communication with any computer is on our fingertips, the concept of Data Security and Privacy has got its importance. Internet Explorer provides a good number of features related to security and privacy. Security features help to prevent people from gaining access to information that one does not want to give others permission to access, such as credit card information. Security features can also protect computers from unsafe downloaded software. Internet Explorer provides security by Security Certificates, which verify the identity of a person or the security of a Web site; by 128-bit secure connection for using secure Web sites; Microsoft Authenticode technology, which verifies the identity of downloaded programs. A certificate is a statement verifying the identity of a person or the security of a Web site. Internet Explorer uses two different types of certificates:

1) A *personal certificate* is verification of user's identity. This information is used when personal information is sent over the Internet to a Web site that requires a certificate verifying the identity. A private key is used for this purpose. When used with e-mail programs, security certificates with private keys are also known as digital IDs.

2) A *Web site certificate* states that a specific Web site is secure and genuine. It ensures that no other Web site can assume the same identity of the original secure site.

A security certificate, whether it is a personal certificate or a Web site certificate, associates an identity with a *public key*. Only the owner of the certificate knows the corresponding *private key*. The private key allows the owner to make a *digital signature* or decrypt information encrypted with the corresponding public key. If some body wants to send the certificate to other people, actually he is distributing a public key, so that they can send encrypted information which is to be decrypted and read with private key. A secure Web site address starts with https (e.g. https://drtc.isibang.ac.in) which automatically sends its

certificate. The digital signature component of a security certificate is an electronic identity card that tells the recipient about the authentication of the source of information and it is not forged or tampered.

Privacy features allow protecting personally identifiable information by helping to understand how Web sites may use this information and by allowing privacy settings that determine whether or not to allow Web sites to save cookies on computer. Privacy features in Internet Explorer include:

- Privacy settings that specify how your computer deals with cookies. Cookies are files created by a Web site that store information on your computer, such as your preferences when visiting that site. Cookies may also store personally identifiable information, such as your name or e-mail address.
- Privacy alerts that let you know when you try to gain access to a site that doesn't meet the criteria in your privacy settings.
- The ability to view a Web site's Platform for Privacy Preferences (P3P) policy.

Protecting Computer from Unsafe Software

Now and then numbers of software are downloaded from different sites. But how good those are for a computer? Whether they have been downloaded from a known, reliable source? To provide a safeguard, Internet Explorer uses Microsoft Authenticode technology to verify the identity of the program. Authenticode technology verifies that the program has a valid certificate: that the identity of the software publisher matches the certificate, and that the certificate is still valid. Note that this does not prevent a poorly written program from being downloaded or run on your computer, but it does reduce the chance of someone misrepresenting a program that is intended to be malicious or intentionally harmful. One can specify different settings for the way Internet Explorer handles downloading programs and files, depending on the zone they are coming from.

b) **Netscape Navigator**

Netscape navigator is the most widely used graphical Web browser and popularly known as Netscape. The salient features of Netscape are:

- it combines text, pictures, graphics and animation in lots of attractive ways;
- it runs reasonably well even on inexpensive dial-up connections;
- consecutive versions are being updated to add new Internet services; and
- its authors give it away for free.

Netscape is the best organised, best performing and most efficient Web tool. It recognises the advanced HTML tags. Netscape navigator supports a host of widely used enhancements to HTML. A clear interface, easy-to-use controls and a straightforward set up also add to this browser's appeal. A version is named as Netscape's Page Composer which has the facilities of both editing the HTML document and browsing it. Thus tedious to and fro movement from HTML Assistant to Netscape can be avoided.

To load a particular URL, the address of the website, just go to the File menu, then Open Location. You'll get a dialog box asking for the address, where you just type in the URL, and then press Go button or press Enter key. The Open button in the toolbar will do the same thing. The grey box at the bottom of the screen tells you where a link will take you, when you move the mouse over it. Links those are already used in the same session turn purple instead of blue. Sometimes if you try to load a page, it will take a very long time. This may be because of the links being down, the machine being slow, the server being overloaded, or a number of other problems. If you get tired of waiting, you can just press the STOP button — it will be red — and the current transfer is stopped. You can tell if Netscape is still trying to download something, because stars will zoom past the big **N** in the upper right corner. Book marking is a way to have Netscape remembers a certain page that you think you'll want to visit again. If you want a hard copy (*a print-out on*

paper) of the current web page, just click on the PRINT button on the toolbar, and then press OK on the next panel that appears. The page will be printed on the printer.

When you are switching from Netscape Navigator or just want to use more than one browser, it's easy to learn and use Internet Explorer. From keyboard shortcuts to bookmarks, you can make the transition simple and fast.

- Keep unique and separate settings for Internet Explorer and Netscape Navigator. If you installed another Web browser after installing Internet Explorer and Internet Tools, some of your Internet Explorer settings may have changed.
- Compare terms in Navigator and Internet Explorer. If you've used Navigator in the past, you may find that words for common actions or features are different in Internet Explorer.
- Keep all of your bookmarks as favourites. Internet Explorer automatically imports all of your bookmarks (without altering your bookmarks as they are displayed in Navigator) as Internet Explorer favourites.
- Use your favourite keyboard shortcuts. Internet Explorer and Navigator share many of the same keyboard shortcuts. In addition, Internet Explorer is fully accessible using the keyboard, so you can park your mouse if you want to.
- Customize how Internet Explorer looks, how you connect, your security settings, and many other options. These settings and more are available by clicking the Tools menu and then clicking Internet Options.

OCR Software

Optical Character Recognition (OCR) is the process of extracting text from an image. This image can result from scanning a paper document or opening an electronic image file. Images do not have editable text characters; they have many tiny dots (pixels) that together form character shapes. These present a picture of the text on a page. OCR analyzes the character shapes in an image and defines solutions to produce editable text. After OCR, you can save the resulting text to a variety of word-processing, desktop publishing or spreadsheet applications. While purchasing scanner, generally an OCR software is supplied free. However, OmniPage Pro is one of well known, powerful and widely used commercial software.

Capabilities of OCR Software

In addition to text recognition, powerful OCR software should retain the following elements of a document through the OCR process.

Graphics: Photos, logos, and drawings are examples of graphics.

Text formatting: Font types, sizes and styles (such as **bold**, *italic* and underlines) are examples of character formatting. Indents, tabs, margins and line spacing are examples of paragraph formatting.

Page formatting: Column structure, table formats, and placement of graphics and headings are examples of page formatting.

The graphics, text and page formatting elements that OmniPage Pro retains are determined by the selected settings done through *Settings Guidelines*. OmniPage Pro only recognizes machine-generated characters such as offset or laser-printed or typewritten text. However, it can retain handwritten text, such as a signature, as a graphic.

Self Check Exercise

18) What do you understand by OCR?

Note: i) Write your answer in the space given below.

ii) Check your answer with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

2.5 MARKUP LANGUAGES

Markup Languages are not programming language. These are used widely for web authoring. Now, what is this markup language? Broadly, markup implies a process involved in combining simple text or information and some special symbols to prepare a meaningful expression to control display. In other words, it is a way to use some special symbols or elements that instruct web browsers how to format the contents of a particular kind of documents, while these special symbols are identified as markup tags. Thus a markup language is meant to give clues about the structure, content, formatting, presentation and behaviour of document either to machine or human beings. They are of two categories — descriptive and procedural markup language. A *descriptive markup language* is to describe the structure and behaviour of a document allowing the author to concentrate on the content and structure. The markup tags are used to describe the structure of the corresponding text which is otherwise known as 'object'. However, it is the responsibility of a browser to render these properly onto a computer screen. But it must be remembered that though the rendering of objects may differ from browser to browser, they have a common significant and meaningful way to display. This category of markup languages gives less emphasis on formatting and presentation. On contrast, procedural markup gives more emphasis on formatting and presentation than its structure, i.e., markup languages which carry the instruction for text processing are known as *procedural markup language*.

SGML and its Origin

Standard Generalized Markup Language (SGML) is a super set of all other markup languages. It is a highly detailed and complex system of presenting information that is designed to allow content to be platform independent and is a powerful complex tool for representing documents of all kinds. It offers the ability to create document specifications of all kinds that can then be used to define and build individua! documents that conform to those specifications. Here, tags decide how a browser can display the contents of that document; how a browser can adopt multimedia representation with graphics, sound, video, animation, etc.; how it can handle and display hypertext links-either within the same file or to other documents or web-accessible services.

SGML was originated with the work begun at IBM in the 1960s to overcome the problems inherent in moving documents across multiple hardware platforms and operating systems. IBM's efforts were named as Generalized Markup Language (GML). GML was originally targeted for local use at IBM, rather than as a generic way of representing documents. This was multi-platform strategy for document preparation-a concept that has become extremely popular today.

Charles Goldfarb, Ed Mosher and Ray Lorie originated GML by the 1970s that a more general version of markup would make documents portable from one system to another. Their work led ultimately to the definition and birth of SGML in the 1980s, which is covered today by the ISO 8879 standard developed in 1986. Several business organizations, commercial and government institutions have adopted SGML. SGML allows users:

- to link files together to form composite document,
- to identify where illustrations are to be incorporated into text files,
- to create different versions of a document in a single file,
- to add editorial comments to a file, and
- to provide information to supporting programs.

Even though SGML is meant for producing documents, the goal is to make them behave more like programs, to behave predictably in a computer-oriented world. Text processing and word processing systems typically require additional information to be interspersed among the natural text of the document being processed. This added information, called 'markup', serves broadly following two purposes: they separate the logical elements of the documents; and they specify the processing functions to be performed on those elements.

Hypertext Markup Language (HTML)

Hypertext Markup Language, coined by Ted Nelson, has got the beautiful combination of two major concepts—Hypertext and Markup Language. It has already been discussed earlier that Markup language is meant for embedding special tags which describe the structure as well as the behaviour of a document. HTML is a markup language that describes the structure of a web document along with a few behavioural characteristics. It is the basic mechanism for distributing information on the web. However, HTML is not a programming language and an HTML document cannot be called a program. Because, a computer program is a set of logically arranged instructions and operations are to be applied to external data used as input. In case of HTML, instruction and data are combined to tell the browser how to render the data contained in the document. HTML ensures plenty of structuring and layout controls to handle the presentation and linkage mechanism necessary to provide hypertext capabilities.

HTML Editors

For authoring web documents some editors are used. FrontPage, Dream Weaver, Hot Metal Pro, HTML Assistant, HTML ed., Internet Assistant, Page mill, Spider, Web Author, etc. are a few to name. However some tools like Netscape Navigator Gold, Netscape's Page Composer, and Microsoft's FrontPage/Internet Explorer have both editing and browsing facilities.

Advantages of HTML

i) HTML is a hyper-media document description language to publish documents on the WWW. HTML documents are written in plain text with the file extension .html or .htm. Exciting facilities of hypertext, the key concept, have added another dimension to HTML.

ii) Ease of creating document: Any editor may be used and the HTML tags may be added in. Any HTML editor that inserts the HTML tags in the appropriate location may be used. Further a conversion utility that takes a document from some other format and converts it to HTML may also be used.

iii) Ease of tracing references: HTML machine has facilities to support for link tracing which means that all references are equally easy to follow forward to their referent, or to backward to their references.

iv) Ease of displaying information: The greatest advantage of HTML is the display of information on any platform, operating system and display cards.

v) Ease of creating new references: HTML facilitates its users to grow their own networks, of simply to annotate someone else's document with a comment (without changing the referenced document).

vi) Ease of information structuring: HTML, by its nature, is helpful in both hierarchical and non-hierarchical organisations of unstructured information. Even the same material can be organised in multiple hierarchical manners.

vii) Global views: Browsers provide table of contents style view supporting easier restructuring of large or complex documents; global and local (node or page) views can be mixed effectively.

viii) Ease of customizing documents: In HTML, text segments can be threaded together in many ways, allowing the same document to serve multiple functions.

ix) Ease of maintaining modularity of information: Using HTML, since the same text segment can be referenced from several places, ideas can be expressed with less overlap and duplication.

x) Ease of maintaining consistency of information: References are embedded in their text, and if the text is moved, even to another document, the HTML uses link information to provide direct access to the references.

xi) Task stacking facilities: The user is supported in having several paths of inquiry activate and displayed on the screen at the same time, such that any given path can be unwound to the original task.

xii) Collaboration: Several authors can collaborate; with the documents and comments about the documents being tightly interwoven (the exploration of this feature has just begun).

Disadvantages of HTML

i) Users Non-interactivity : The web pages created using HTML are static and do not provide user interactivity.

ii) Disorientation: The tendency to lose one's sense of location and direction in a nonlinear document.

iii) Cognitive overhead: Additional effort and concentration are necessary to maintain several tasks or trails at one time. These problems may be at least partially resolvable through improvements in performance and interface design of HTML systems, and through research on information filtering.

eXtensible Markup Language (XML)

HTML is more like a formatting language. Thus it is difficult to pull out what kind of data is stored inside the HTML document. To solve this problem, a fixed structure of the document is required so that everybody may write the document in the same structure. Thus if the information/document is transferred from one system to another it will not be a problem as the structure of document is always same. The definition of the structure of

document is known as Document Type Definition (DTD). To prepare DTD, domain specific tags are needed. For information interchange, which is not possible with HTML. Hence, the e**Xt**ensible **M**arkup **L**anguage (XML) was developed. According to the abstract from the XML Specification version 1: *"The extensible Markup Language (XML) is a subset of SGML. Its goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML. XML has been designed for ease of implementation and for interoperability with both SGML and HTML."*

Basically, XML is used to describe data and to focus on 'what data is', but HTML is to display data and to focus on 'how data looks'. HTML is about displaying information, whereas XML is about describing information. XML is useful to carry data. It is a much like HTML. Tags are not predefined in XML and one's own domain specific tags must be defined. XML uses a DTD (Document Type Definition) to describe the data; hence XML with a DTD is designed to be self-descriptive. XML is not a replacement for HTML. XML is used to describe the data, while HTML is used to format and display the same data, using CSS (Cascading Style Sheets) or XSL (eXtensible Style-sheet Language). However, XML is still under development and needs more refinement.

XML Editor

XML Spy Suite is a comprehensive and easy-to-use product family that facilitates all aspects of XML Application Development. The product family consists of the XML Spy Document Framework and XML Spy IDE. XML Spy Document Framework consists of XSLT Designer and XML Spy Document Editor. Help on each member of the product family is available using the Help menu option. XSLT Designer is a new approach to automate the writing of complex XSLT Stylesheets using an intuitive drag-and-drop user interface. XSLT Designer creates advanced electronic forms for use with XML Spy Document Editor which is a word processor type editor, supporting electronic form-based data input, graphical elements, tables, as well as real-time validation using XML Schema. XML Spy IDE is a solution for XML-based application development, allowing easy creation and management of XML documents, XML schemas, as well as XSLT Stylesheets. XML Spy Document Editor enables you to edit XML documents based on templates created in XSLT Designer. XML Spy Document Editor is available in three versions: As a stand-alone application, Integrated as a separate view within the XML Spy IDE user-interface, and as a Browser Plug-In for Internet Explorer.

XML Document Editor features include free-flow WYSIWYG text editing, form-based data input, presentation and editing of repeating XML elements as tables, real-time validation and consistency checking using XML Schema.

Self Check Exercises

19) What are the parts of a HTML document?

20) What do you understand by XML?

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

2.6 SUMMARY

The term 'software' denotes the programs or predefined sequences of instructions that a computer follows while executing a task. The software is an intellectual product and is neither physically visible nor tangible. The two principal types of software are: System software and Application software. System software helps the computer to manage its internal resources as well as enables the application software to interact with the computer. Application software performs useful work on general basis. Custom-written software is designed for a particular customer in mind. It is meant for highly specialised tasks. Packaged software, also known as software package, is a package of programs available separately for use by the computer users. It has to be loaded on to the computer at the time of requirement. System software controls the systems programs, manages the storage and processing resources, provides support to the system's routine functions, and assists in developing system tools like compilers, interpreters, assemblers, etc. On the other hand application software are programs written by keeping a particular job in mind. The programs can be used for general-purpose applications or specific purpose applications.

System software and application software act as interfaces between users and computer hardware. The system software controls the execution of application software and provides other data functions, such as data storage. For example, when the electronic spread-sheets (i.e., application software) is used in an IBM compatible PC, the storage of the worksheet files on disk is handled by MS-DOS, the computer's operating system (i.e., System software).

Earlier use of computers was confined to skilled professionals like analysts and programmers. The work was confined only to a separate section of the organization, i.e., the computer section. The introduction of microcomputers (PCs) has dramatically shifted its usefulness and became user-friendlier. Now-a-days, PCs are used by persons from all walks of life and in all areas of knowledge. This is mainly due to the development of soft wares/software packages fitting to the needs of personal computers. The PC software packages are now used by all kinds of users. Thousands of software packages are commercially available which can be run on PCs, but the most popular ones are those, which belong to the productivity software packages.

Computers and communications are becoming the integral parts of the modern lives. The use of computers has become closer to man's daily living. It is due to the advancement of the software technology. Softwares are being developed in all sizes, sophistication, to fit into the requirements of all walks of life and to be used in all the areas of social living, In earlier days, the softwares were developed in isolated manner, but today, they are being developed jointly by different organisations, each organisation shouldering the responsibility of developing a particular component of the software. Software products and systems are becoming larger, more complex and more sophisticated than the earlier ones.

As the software development goes by adding new features onto them, there is a corresponding development in hardware technology to cope with these new features of softwares. The computer units are produced with faster and powerful processors, more RAM capacity, and more disk space and so on to make them compatible to the new software products.

Any system which allows the fast, cost effective production of documents using high quality publication techniques must be of benefit in facilitating information dissemination. DTP is a fully-fledged and widely used in the library and information centre environment increasing library professional's awareness in respect of its nature and potential. The use of DTP by library and information professionals are gradually varied and not confined to what is conventionally regarded as the province of publishing, such as the production of newsletters, bulletins or booklets. It is also useful as resources that library can provide for its users in the form of printed material. In case of special library attached to the research institution, the scientists who use the library to keep track of the work of their colleagues are at the same time and as part of the same process, involved in communicating their

own results to others. The concerned library personnel can render effective services towards its user in this regard if they are well trained for handling desktop publishing techniques.

2.7 ANSWERS TO SELF CHECK EXERCISES

1) Software is a set of programs and accompanying documentation that direct computers to perform desired job. In simple terms, a software program is a set of instructions for a computer. In other words, it refers to all programs that are written to be executed on a computer.

2) The main function of software is to make use of hardware to perform a desired job. For carrying out different jobs, different types of software are designed and developed.

3) System software is a set of one or more programs, designed to control the operation of a computer. System software makes the operation of the computer system more effective and efficient. It provides environment to develop and run application software.

4) Operating systems can be divided into the following types: Batch Processing System, Multitasking System, Time Sharing, and Multiprocessing System.

5) The commands are grouped into two: internal and external commands. Internal commands (viz. DIR, DEL, COPY, etc) are within a single separate file COMMAND.COM, called command processor and are loaded into the RAM at the time of booting. Hence they are executed directly from RAM. External commands (viz. CHKDSK, EDIT, FORMAT, etc) are the names of the programs stored as an executable disk file. So they must be loaded into the RAM when needed.

6) A window is a rectangle that appears on the screen and displays information for some specified purpose. The operating system (or the operating environment) which is window based can display several windows simultaneously on the screen each meant for different purpose and may belong to different programs like word processing, spreadsheets, etc. However, conventionally Windows is an operating environment developed by Microsoft that lays a graphical user interface sheet around the MS-DOS or PC-DOS operating system. Windows makes use of windows extensively for its functioning.

7) UNIX consists of three layers - the UNIX Kernel, the UNIX Shell and UNIX utilities/Application layer. These layers are developed from protection point of view and each layer is meant for specific functions. *The UNIX Kernel*: This is the most important and the innermost layer. It has a process scheduler, a hierarchical file structure and mechanisms for processes to communicate with each other. *The UNIX Shell*: This is the middle layer which contains command interpreter. It executes users commands and programs. *UNIX Utilities/Applications layer*: This is the outermost layer containing programming tools and application programs. The programming tools include assemblers, compilers, editors, etc. The application programs include packages for database management, word-processing, graphics, etc.

8) Linux kernel offers the common features like *Application support, Graphical User Interface (GUI), Hardware support, Memory Management, Multitasking, Multi-user, Network client/server, Networking connectivity, Open Source.*

9) Generally, no source code used to be distributed earlier with the software. But with the open source movement, now source codes are supplied with the software so that any interested person can add some new features to that software. Linux is

open source software that means all source code is available, including the whole kernel and all drivers, the development tools and all user programs; also, all of it is freely distributable. Plenty of commercial programs are being provided for Linux.

10) The Network Operating System is the main system software to work in a network environment. And most of the recent computers have the facility for network connections. A network is a group of computers connected to each other or to a central server so that they can share resources such as documents and printers. A network can have a variety of platforms having different systems. In addition, a server can act as a client to another server. A machine designated as a client, however, cannot be a server. There are many operating systems that can be used as NOS. It also helps to gain access to the Internet; and send and receive messages by using e-mail or help in connecting office and home computers. This is accomplished by the Network Neighborhood icon which appears on the desktop. By double-clicking this icon, one can get an overview of all the computers within the workgroup or any Netware servers to which the computer is connected.

11) Interpreter is one which translates one statement of source program at a time into one or more machine instructions (i.e., object code) subject to prompting programmer to debug, if any, the source program and then instructs the computer to perform immediately those instructions before the control moves towards the next statement.

12) Compiler is one which is used to store the high level language program, scan it and translate the whole program at a time into an equivalent machine language program that is a compiler translates the source program into a complete machine language program/object program.

13) A word processor is the software that performs the word processing task. Word processors facilitate faster production of documents as compared to traditional typewriters. Although the textual matter is typed into the computer through the keyboard, which resembles the traditional typewriter, the keyboard has certain special (functional) keys which the typewriter does not have. The task of editing the textual matter; adding and deleting of sentences, paragraphs; setting of margins; spelling checking, etc. are much easier and almost simultaneous in word processing rather than typewriting. Therefore, long documents, repetitive letters, memos, and general correspondence are the most widely used applications for word processors. The most significant advantage of word processors is increased productivity.

14) The advantages of DTP are: a) Low Cost Publication, b) Minimum Time Requirement, c) Flexibility in Qualities, d) Convenient for Designers, e) System Integration In Office.

15) Creating a presentation in Microsoft PowerPoint involves starting with a basic design; adding new slides and content; choosing layouts; modifying slide design, if you want, by changing the colour scheme or applying different design templates; and creating effects such as animated slide transitions. The *New Presentation* task pane in PowerPoint gives a range of ways with which to start creating a presentation. These include:

    i) Blank Presentation: In this case, slide starts with minimal design and no colour. But presentations with colour scheme, fonts, and other design features with a wide range of predefined slide formats and Clip Art graphics libraries can be implemented. These predefined slide formats help to create slides quickly based on standard layouts and attributes applied to them.

    ii) Existing Presentation: New presentation is prepared based on already written and designed one. This command creates a copy of an existing presentation so the design or content may be changed as required for the new presentation.

iii) Design Template: Presentation is made based on a PowerPoint template that already has a design concept, fonts, and colour scheme. Some of the templates come with PowerPoint, and some can be borrowed from the web site also. The Auto-Content Wizard is used to apply a design template that includes suggestions for text on slides. Microsoft.com helps to select templates from the Template Gallery which is arranged according to type of presentation.

iv) Other sources: Slides content from the other sources of presentation or text (e.g. MS Word) can also be inserted. OLE facilities are also available.

16) The PowerPoint allows the Object Linking and Embedding (OLE) facilities to the presentation. The main differences between linked objects and embedded objects are where the data is stored and how the data is updated after it is placed in the destination file. When an object is linked, information is updated only if you modify the source file. Linked data is stored in the source file. The destination file stores only the location of the source file and displays a representation of the linked data. When an object is embedded, information in the destination file doesn't change if you modify the source file. Embedded objects become part of the destination file and, once inserted, are no longer part of the source file.

17) The main functions of electronic spreadsheets are:

- To help in tabular presentation of any quantitative data for computation,
- To help in recalculation of data easily reflecting changes, if any, in the existing data set,
- To perform simple and statistical calculations,
- To help in determining relationships, if any, among data sets,
- To help in creating budgets, financial reports, forecast cash needs,
- To facilitate graphical presentations.

18) Optical Character Recognition (OCR) is the process of extracting text from an image. This image can result from scanning a paper document or opening an electronic image file. Images do not have editable text characters; they have many tiny dots (pixels) that together form character shapes. These present a picture of the text on a page. During OCR analyzes the character shapes in an image and defines solutions to produce editable text. After OCR, you can save the resulting text to a variety of word-processing, desktop publishing or spreadsheet applications.

19) HTML document has the following parts:

- A *head*, which identifies the document as HTML and establishes its title;
- A *body*, which contains the content for a web page. This is where all displayed text on a page comes from, as well as links to graphics, multimedia information, and to locations inside this HTML file or to other web documents; and
- A *footer*, which labels the page by identifying its author, date of creation and version number.

But HTML has flexibility of skipping some of these elements. For example, HTML does not include a separate tag to denote page footer.

20) The e**Xt**ensible **M**arkup **L**anguage (XML) was developed. According to the

abstract from the XML Specification version 1: "The extensible Markup Language (XML) is a subset of SGML. Its goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML. XML has been designed for ease of implementation and for interoperability with both SGML and HTML." Basically, XML is used to describe data and to focus on 'what data is', but HTML is to display data and to focus on 'how data looks'. HTML is about displaying information, whereas XML is about describing information. XML is useful to carry data. It is a much like HTML. Tags are not predefined in XML and one's own domain specific tags must be defined. XML uses a DTD (Document Type Definition) to describe the data; hence XML with a DTD is designed to be self-descriptive. XML is not a replacement for HTML. XML is used to describe the data, while HTML is used to format and display the same data, using CSS (Cascading Style Sheets) or XSL (eXtensible Style-sheet Language).

2.8 KEYWORDS

Buttons : A button is a kind of icon which is stimulated on-screen and is activated by a mouse or other pointing devices to execute a command.

Clicking : Clicking refers to pointing an item on the screen and then pressing and releasing mouse button.

Desktop : Windows desktop is the basic windows working area. As the system is started, the first thing one can see is the desktop, the area on the screen where one works. The following components appear immediately on the desktop.

Dialog box : A dialog box is a box that appears on the screen and displays an option among the alternatives, for example 'Y' for yes and 'N' for no or typing the name of a file.

Dragging : Dragging refers to holding down the left mouse button while sliding the mouse to a new position on the screen, and then releasing the button.

Icons : An icon is a small picture or graphical display used to represent a command, a program or a task. It has to be activated by a mouse or other pointing devices. In the figure, the desktop contains different icons.

Internet Explorer : It enables user to connect the Internet.

Microsoft Outlook Express : It is a component of Internet Explorer that integrates an e-mail and newsgroup services. It has the address book to store and retrieve e-mail addresses.

Mouse Movements : The mouse points to the location of the desired item on the screen with the help of a cursor called mouse pointer.

My Computer : My Computer gives a quick overview of all the files and programs available in the system. Any program installed shall get into 'My computer' folder automatically.

My Documents : It gives a quick overview of all the documents saved in the system. Any document that is saved gets into My Documents folder automatically. Besides these basic components, many other icons can be kept on the desktop.

Network Neighborhood : It helps to get a glimpse of what is available with other computers on the network.

Pointing : Pointing refers to the positioning of the mouse pointer at a specific location on the screen.

Recycle Bin : It keeps all deleted files, to retrieve them later.

Release : A minor upgrade in a software product. It is usually indicated by a change in number after the decimal point such as 3.0, 3.1, 3.11, 3.2, etc.

Use of year : Occasionally software developers prefer to add the year of release to a new version rather than a numeral for their products. For example, Windows 95 (but not, Windows 4.0), Windows 98, Windows 2000, etc.

Version : A major upgrade in a software product. It is usually indicated by numbers such as 2.0, 3.0, etc.

2.9 REFERENCES AND FURTHER READING

Aruna, A (1995). *Role of desktop publishing (DTP) in information centre*. In: *Current Research in Library and Information Centre*. Chapter M. pp 164-76.

Beaumont, J (June 1965). Desktop publishing: what does it mean to libraries? *Canadian Library Journal*. 45(3), 41-44.

Borrell, J (December, 1985). Electronics publishing lands on the desktop. *Minimicro Systems*. pp 85-92.

Davis, D M (1988). *Computer aided composition*. In: *Encyclopedia of micro computer* V3, Compuserve to computer programme: outlines. New York: Marcel Dekker. pp 7-21.

December, J and Ginsburg, M (1996). *HTML 3.2 and CGI unleashed*. Professional Reference Edition. Indianapolis: Samnet Publishing.

December, John and Randall, Neil (1996). *The World Wide Web unleashed*. Indianapolis: Sammet Publishing.

Devika, P. Madalli (Jan, 1998). *World Wide Web: hotlink to information*. In: Practical *orientation to Internet*, edited by Devika, P. Madalli (Course material of the DRTD Workshop, 28-30 Jan, 1998, Bangalore). Bangalore, DRTC and NISSAT. Paper BA (pp 1-18).

Edkins, L (April, 1990). Library application of desktop publishing in practice: a tool and a resource. *Program*, 24(2), 155-168.

Gupta, Vikas (1999). *Rapidex computer course*. 4th ed. New Delhi: Pustak Mahal.

Jacob, Nielsen (1990). *Hypertext and hypermedia*. Boston: Academic Press.

Jain, Satish (1999). *Information technology*. New Delhi: BPB Publications.

Jaiswal, S. (2000). *Information technology today*. New Delhi: Galgotia Pub.

Mahapatra, M and Ramesh D.B. (eds.) (2004), *Information technology application in libraries: a textbook for beginners*. Bhubaneswar: Reproprint.

Panigrahi, Pijushkanti (2004). *Markup languages in web authoring*. In: *Information technology application in libraries: a textbook for beginners*, edited by M. Mahapatra and D.B. Ramesh. Bhubaneswar: Reproprint. pp 496-509

Panigrahi, Pijushkanti and Prasad, A.R.D. (1997). Data recovery utility for librarians. *Vidyasagar University Journal of Library and Information Science*, 2, 32-40.

Prasad, A.R.D. (1996). *Some reflection on the impact of IT on LIS profession*. In: *Advances in information technology: impact on library and information field*, edited by I. K. Ravichandra Rao and A.R.D. Prasad. (Proceedings of the 30th DRTC Workshop, 28-30 Oct, 1996, Banagalore). Bangalore: Documentation Research and Training Centre. Paper AC.

Prasad, A.R.D. (1998). *Browsers: the interface to the web*. In: *Practical orientation to Internet*, edited by Devika P. Madalli (Course material of the 31s1 DRTC Workshop, 28-30 Jan, 1998, Bangalore). Bangalore: DRTC and NISSAT. Paper BB (pp 1-8).

Townsend, K and Tapehouse, K (1987). Word processing and desktop publishing. *Communication Technology Impact*, 9(7), 1-4.

Tuck, B (1989). Desktop publishing: what it is and what it can do for you. *Aslib Proceedings*, 41(1), 29-37.

West, S (1988). Desktop publishing applications for libraries. *Bulletin of the Florida Chapter of the Special Libraries Association*, 20(2), 28-33.

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 3 PROGRAMMING

Structure

3.0 Objectives

3.1 Introduction

3.2 Algorithm and Flow Charts

3.3 Program Development Life Cycle

- 3.3.1 Program Specification
- 3.3.2 Program Design
- 3.3.3 Program Code
- 3.3.4 Program Testing
- 3.3.5 Program Documentation
- 3.3.6 Program Maintenance

3.4 Types of Programming

- 3.4.1 Procedural Languages
- 3.4.2 Object Oriented Programming Languages
- 3.4.3 Special Purpose Languages

3.5 Generations of Programming Languages

- 3.5.1 Machine Language
- 3.5.2 Assembly Language
- 3.5.3 High Level Languages
- 3.5.4 Fourth Generation Languages
- 3.5.5 Natural Languages

3.6 Summary

3.7 Answers of Self Check Exercises

3.8 Keywords

3.9 References and Further Reading

3.0 OBJECTIVES

In the previous Unit you have learnt about basics of computer software. In this Unit we will discuss about programming and programming languages.

After reading the Unit, you will be able to:

- understand the basics of programming;
- be acquainted with different categories of programming languages; and
- become familiar with types programming languages.

ignou THE PEOPLE'S UNIVERSITY

3.1 INTRODUCTION

A computer can neither think nor make decisions on its own. To solve a problem using computers, a set of instructions should be given to the computer in a form which it can understand. The process of writing these instructions is called programming.

Understanding what programming is can be explained by taking a simple and well-known example. Consider searching a dictionary for a word.

The manual solution to this problem is to take out the dictionary and search pages for the initial letter of the word. Once the word is found, search for the subsequent letters of the word and on finding the word, read out its meaning given in the dictionary.

Let us assume a dictionary is available in a computer. We need to write a program to solve this problem of searching a word in the dictionary. The algorithm to solve the problem is given below:

1) Take the word to be searched as input

2) Match the words available in the dictionary one by one until the exact match is found. (It is the simplest logic but we can apply more efficient logic)

3) Display the meaning as output for the input word.

This solution is in natural language which computers cannot understand. Therefore, we need to write it in a language, as a set of instructions, which computers can understand. Such a set of instructions that are arranged in a sequence that guides the computer to solve a problem is called a program.

The process of writing a program is called **Programming**. It involves *solving a problem and writing it in a form which a computer system can understand and execute.* While considering how to solve a problem, there are three main specifications which should be taken into consideration:

- What information flows into the system?
- What information flows out of the system?
- What the system does with the information?

Computers cannot understand natural languages such as English or Hindi for solving problems. To communicate instructions and commands we need programming languages. And learning programming languages requires learning the symbols, words and rules of the language. There are many languages available to do programming. These are developed for specific applications. Based on these applications, programs are of two types, *system programs* and *application programs.*

System program: These are the programs which make computers available to use. Operating Systems such as DOS and Windows which consist of many other programs for controlling input/output devices, memory, processor etc are system programs. The *C language* is widely used to develop system software.

Application program: Application programs are designed for specific application such as library automation, information retrieval, information processing, inventory control, etc. These applications should be able to get input, produce output, do calculations, store and retrieve data. There are many languages which contain instructions to support all these operations. Before continuing with programming, we look at what algorithms and flow charts are and how these helps in programming.

Self Check Exercise

1) What do you call a list of instructions for the computer to follow to process data?

Note: i) Write your answer in the space given below.

ii) Check your answer at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

3.2 ALGORITHM AND FLOW CHARTS

Computers are used for problem solving, automating and processing of data and information. The computer does it based on the instructions it receives through various programming languages such as C, C++ or a variety of others. The set of instructions given to the computer to perform a job is called a program which is nothing but an algorithm expressed in language known as programming language. Thus, as mentioned earlier writing a program requires knowledge of the programming language. However, the set of instructions can also be expressed independent of any programming language which is called *algorithm*. There are many definitions of algorithm. According to Dromey, "*algorithm* consists of a set of explicit and unambiguous finite steps which, when carried out for a given set of initial conditions, produce the corresponding output and terminate in a finite time."

In the previous section, we discussed about an algorithmic process of looking for the meaning of a word from a dictionary. This kind of algorithmic process can be used for problem solving using computers.

To obtain solution to a problem through the computer, usually we have to provide the computer program with input or data. The program then takes this input and manipulates it according to its instructions and eventually produces an output which represents the computer solution to the problem.

Algorithms are written in a unique language and style. The way and language used to write an algorithm is its representation. A natural language like English can be used to write algorithms. Let us take a simple problem of adding numbers from 1 to 'n' (that is if 'n'= 4, then it should add 1+2+3+4). If the algorithm is written for this problem in natural language, it would read as follows:

> Initially set the sum value to 0 and count to 1. Accept 'n' as input from the user. If the value of 'n' is 0 then print the value of sum as 0. If the value is greater than 0 then add the count till it reaches the input value 'n'. Each time after addition increase the count value.

It can be noticed that even for this small problem, the natural language is unstructured. In case of huge problem, the natural language algorithm will be difficult to understand while programming.

The next choice of writing the algorithm may be in formal languages like C, C++, etc. But, at the initial phase of design one would be thinking and writing at a highly abstract level. Using a programming language forces the use of punctuations and syntax properly which cannot be done at this early stage of design. The algorithm should be syntax free of the programming language.

A language between a natural and formal language is necessary to write algorithms efficiently. There is a notation called *pseudo code* which is most commonly used to write

algorithms. It is simple, readable and has no grammatical rules. As it has very well defined structure it is easy to visualize the organisation of statements. As pseudo code resembles many of the programming languages, it is easy to convert the algorithm to programs.

Algorithm for the above problem in pseudo code is as follows:

Step 1 Initialize the variable Count to 1 for counting the addition.

Step 2 Initialize the variable Sum to 0 to store the sum.

Step 3 Get the input value 'Num' for number of times sum has to be performed.

Step 4 Add Count with Sum and assign it to Sum.

Step 5 Increase Count to next value.

Step 6 Repeat steps 4 and 5 till Count exceeds the value of Num.

Algorithms can also be pictorially represented using **flow charts**. The flow charts give easy and clear understanding of algorithms. The understanding of algorithm is made easy by flow charts, compared to textual representation algorithm.

Flow chart symbols

Flow chart has many symbols for basic operation in programming. The flow chart symbols are:

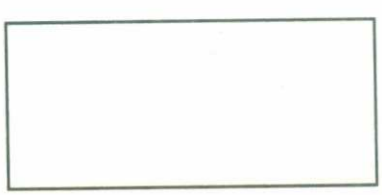
i) *Terminal*

This is usually to represent terminal points in a program like BEGIN, END, START, STOP.

ii) *Input/Output*

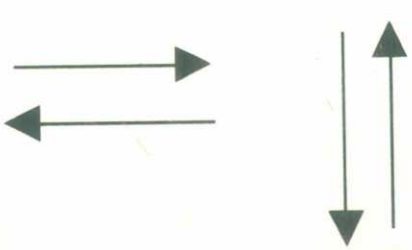
The parallelogram represents the Input/Output function, i.e., making the data available for processing (input) or recording of the processed information (output). This step implies obtaining a number from an input device (say, the keyboard and storing it in the storage location named 'A')

iii) *Process*



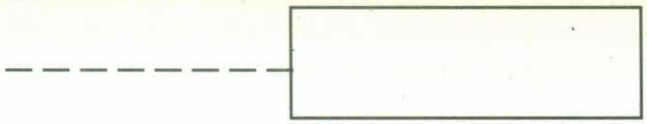
The rectangle represents the processing operation. A process changes or moves data. An assignment is normally represented by this symbol.

iv) *Flow direction*



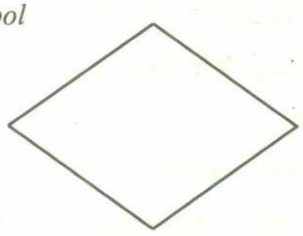
Lines or arrows represent the flow direction – the flow of control. Normally the flow direction is from left to right or top to bottom.

v) *Annotation*



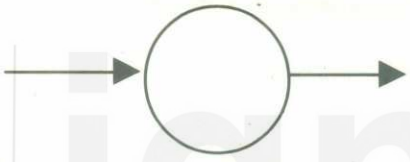
A broken line and rectangle as above represent the annotation function – the addition of descriptive comments or explanatory notes for clarification of some statements.

vi) *Decision making symbol*



The diamond represents a decision or switching type of operations that determines which of the alternative paths is to be followed. A decision indicates a question that can be answered *yes* or *no (*or *true* or *false)*.

vii) *Connector*



Circle is connector symbol which is used to represent functions in a flow line, i.e., the part of flow chart which continues in the next or another page.

viii) *Pre-defined process*

The double sided rectangle represents a named process that consists of one or more operations or programming steps that are specified elsewhere such as a module or subroutine.

Self Check Exercise

2) What is the difference between an algorithm and flow chart?

Note: i) Write your answer in the space given below.

ii) Check your answer with the answers given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

3.3 PROGRAM DEVELOPMENT LIFE CYCLE

Programming is a six-step procedure for creating the list of instructions. Normally, one would assume programming involves typing or keying in statements into a computer. However, this is only one of the steps in program development. The six steps are as follows:

3.3.1 Program Specification

Program specification is also called program definition or program analysis. The step involves specifying five tasks, namely (1) the program's objectives, (2) the desired output, (3) the input data required, (4) the processing requirements, and (5) the documentation. Desired output is totally as per the users requirements which is achieved out of requirement analysis by the software system analyst.

Program objectives

We have learnt that programming involves solving problems. Therefore, at the outset, it is essential to make a clear statement of the problem that you are tying to solve. An example could be, "To develop a bibliographic database that can be searched by various fields".

Desired outputs

It is best always to specify outputs before inputs. The end result that is required from the computer system should be noted down.

Input data

Once the output is known, the input data and the source data can be determined and identified.

Processing requirements

In this step, the processing tasks that must happen for input data to be passed out into output is defined.

Program specification document

Ongoing documentation is essential in programming. All the activities such as the program objectives, desired outputs, needed inputs and required processing should be recorded.

3.3.2 Program Design

After program specification, the program design is carried out. Here, you plan a solution using structured programming techniques. These techniques consist of top-down program design, pseudocode, flowcharts and logic structures.

Now that outputs and inputs of the program are defined, a top-down program design can be used to identify the program's processing steps. Such steps are called program modules (or just modules). Each module is made up of logically related program statements. Pseudocode is an outline of the logic of the program you will write. It is like doing a summary of the program before it is written. Flow charts have already been discussed in the previous section. Finally logic structure involves defining the sequence (one program statement follows the other), selection (when a decision must be made) and loop (describes a process that may be repeated as long as a certain condition remains true).

3.3.3 Program Code

Writing the program is called coding. Here the logic that was developed in the program design step is used to actually write the program. The program code instructs the computer what to do. Coding is what generally people think of when they think of programming. As pointed out earlier, however, it is only of the six steps in the programming processing.

3.3.4 Program Testing

Debugging is a programmer's word for testing and then eliminating the errors. It means running the program on a computer and then fixing the parts that do not work. Programming errors are of two types, syntax errors and logic errors.

A syntax error is a violation of the rules of the programming language. For example, in C++, each statement must end with a semicolon (;). If the semi-colon is omitted, the program will not run due to a syntax error. A logic error occurs when the programmer uses an incorrect calculation or leaves out a programming procedure. For example, a payroll program that does not compute overtime hours would have logic error.

Several methods have been devised for finding and removing both syntax and logic errors. These are:

- Desk checking: In desk checking, a programmer proofreads a printout of the program line-by-line looking for the errors.
- Manually testing with sample data: Using a calculator and sample data, a programmer follows each program statement and performs every calculation. Looking for programming logic errors, the programmer compares the manually calculated values to the correct values.
- Attempt at translation: The program is run through a computer, using a translator program. The translator attempts to translate the written program from the programming language into the machine language. Before the program will run, it must be free of syntax errors. Such errors will be identified by the translating program.
- Testing sample data on the computer: After all syntax errors have been corrected, the program is tested for logic errors. Sample data is used to test the correct execution of each program statement.
- Testing by a select group of potential users: This is sometimes called *beta testing*. It is usually the final step in testing a program. Potential users try out the program and provide feedback.

3.3.5 Program Documentation

Documentation consists of written descriptions and procedures about a program and how to use it. It is not something done just at the end of the programming process. Program documentation is carried on throughout all the programming steps. This documentation is typically within the program itself and is in printed documents. In this step, all the prior documentation is reviewed and finalized. Documentation is important for users, operators and programmer who may be involved with the program in the future.

3.3.6 Program Maintenance

The final step is program maintenance. The purpose of maintenance is to ensure that current programs are operating error free, efficiently, and effectively. Activities in this area fall into two categories: operations and changing needs.

Operations activities concern locating and correcting operational errors, making programs easier to use, and standardising software using structured programming techniques. For properly designed programs these activities should be minimal. Changing needs are unavoidable. All organisations change over time, and their programs must change with them.

3) What are the six steps in programming? What role does coding play in program development?

Note: i) Write your answer in the space given below.

ii) Check your answer with the answers given at the end of this unit.

.....

.....

.....

.....

.....

.....

.....

.....

3.4 TYPES OF PROGRAMMING

We have learnt in the previous sections that programming is writing a set of instructions. If it as simple as that then why can't we use one programming language? Programming languages are designed to meet specific needs. Consequently, one language may be better suited than others for writing certain kinds of functions. There are programming languages that are used for solving engineering problems, writing out complicated sales reports, interact with databases, manipulate graphics or even design web pages. So, with proliferation of programming tasks, proliferation of programming languages has also taken place. Let us see the various families of programming languages and the types of programming languages that belong to them.

3.4.1 Procedural Languages

Languages in which programs are written as sequence of statements that manipulate data item and change the contents of memory cells are called procedural languages (also called imperative languages).

The fundamental operation of these languages are storing and retrieving data values. For example;

a = 1

This statement stores value 1 in location 'a'

c = a + b

This statement retrieves '*a*' and '*b*', adds and stores the result in '*c*'. Some of the procedural languages are discussed below.

FORTRAN

The name FORTRAN derives from *FORmula TRANslation*. The name itself indicates that 'application of language with 'formula' of engineering-type application'. The first commercial version of FORTRAN was released in 1957. This makes FORTRAN the first high-level programming language. FORTRAN has some features ideally suitable for problems that are heavily mathematical or computational oriented. High-level languages are basically English like computer languages where we have 'higher' level of understanding

as compared to Assembly and Machine languages. Programming languages like COBOL, PASCAL, C, C++, BASIC etc also come under this category.

COBOL

The name COBOL derives from *Common Business Oriented Language*. COBOL was designed to serve business needs such as managing inventories and payrolls. In such applications summary reports are important output. Also data file handling is extensively involved. Much of the business world concerns updating master files with change in transaction files. For example master file contain names, manufacturers, and quantities available for various items inventory; a transaction file would contain names, quantities and items sold out of inventory over a period of time. The master file would be updated from the transaction file on weekly or daily basis to reflect new quantities available and print summary report.

PASCAL

The programming language Pascal was named after Blaise Pascal, the inventor of Pascaline calculator.

The design of Pascal was to easily learn and enforce good programming techniques. Pascal looks very similar to pseudocode. So it is easy to read and the syntax is easy to learn. Even though Pascal is not much useful as commercial language, but programming environment like Delphi uses Pascal language with a facility to develop windows based applications with modern graphical user interface.

C

C was developed in early 1970s by Dennis Ritchie at AT&T Bell Laboratories. It was originally developed for system programming, particularly for writing operating system UNIX. It is a popular general-purpose language for two reasons. One is the relationship with UNIX. Second reason is its efficiency that is, speed with which its operations can be executed. The efficiency derives from the fact that C programs can make use of low-level information such as knowledge of where data is stored in memory, yet has powerful statements and portability to machines that high-level languages offer.

The strong feature of C is that it provides a data type called pointer. Variables of pointer type contain memory addresses instead of integers, real numbers and characters. For example the statement,

```
int* intPointer;
```

declares *intPointer* as a pointer variable that will contain the address of a memory location containing integer data.

```
int A = 3;
intPointer = & A.
```

The first statement declares a variable called 'A' and assigns the value '3' to that. The second statement makes the intPointer pointing to 'A'. We can also assign the integer value '10' to variable 'A' by the statement

```
*intPointer = 10;
```

The powerful use of pointers are in the areas of writing system programming, operating systems, assemblers, and programs that allows the computer to interact with input/output devices and so on.

For example consider a problem of writing a device driver (software to interact with a hardware device) for mouse on a PC. The port in which the mouse is connected reads the changes in mouse position by voltage levels. It stores voltage levels in a fixed location of memory. C provides facilities to access this memory and manipulate it.

Even though C is powerful for system programming it is widely used for general purpose programming also.

Ada development was for a common high-level language needed for various branches of United States armed services for defense contracts. Ada is a large language like C++ and has been accepted not only in the defense industry, but also for other technical applications and is a general-purpose language as well.

Ada is known for its multi-processing capability – the ability to allow multiple tasks to execute independently and then synchronize the communication when directed.

3.4.2 Object Oriented Programming Languages

Object Oriented Programming started by Alan Kay's work at Xerox Palo Alto Research Centre in early 1970s. The resulted language is Smalltalk.

The concept of Object-oriented programming is dividing the major task into sub tasks. The program can be thought of a giant statement executor designed to carry out major tasks, even though the main program may simply call the various modules that do the sub task work.

To understand objects, let us take an example of library system. Circulation, Acquisition, Catalogue, etc are major objects. Each object is an example of a task performed in library system.

JAVA

Java programming is associated with the World Wide Web. Java programs can access data across the Web as easily as they access data from a local system. According to the Sun Microsystems, Java is a simple, object oriented, distributed, interpreted, robust, secure, architecture neutral, portable, high-performance, multithreaded and dynamic language. Multithreading is the ability of an application to perform multiple tasks at the same time. Multithreaded programs can be created through Java.

3.4.3 Special Purpose Languages

The procedural languages that we saw so far are more or less general-purpose languages. There are several special purpose languages designed for specialised tasks. Here we take three representative special purpose languages which are very popularly used.

Structured Query Language (SQL)

SQL is designed to be used with databases, which are collections of related facts and information. A database stores data centrally. The user of the database must be able to add new data and to retrieve data already stored. For example, a database of library catalogue may contain titles of book, its author's name, publisher name, year of publication, etc. The database of catalogue user should be able to add information on a book and to retrieve information on a book already in the database. In addition, databases can also be queried with the user posing questions to the database. For example, the library catalogue could be queried to reveal books by a particular author or by a publisher. Such queries may be framed in SQL. SQL is the language used to frame database queries. Sample SQL statement is as follows:

```
SELECT BOOK_TITLE, PUBLISHER, YEAR

FROM CATALOGUE

WHERE AUTHOR NAME = "GUHA";
```

This SQL statement will retrieve all the book titles with publisher and year details of the author Guha.

Practical Extraction and Report Language (PERL)

Perl is designed to scan arbitrary text files, extract various kinds of information that is contained within the text, and print reports based on the extracted information. The language syntax is somewhat based on C. Perl uses sophisticated pattern matching techniques to speed up the process of scanning large amounts of data for a particular text string.

Hyper Text Markup Language (HTML)

This is the language used to create HTML documents that, when viewed with Web browser software, become Web pages. An HTML document consists of the text to be displayed on the Web page, together with a number of special characters called tags that achieve formatting, special effects, and references to other HTML documents. Tags are enclosed in angle brackets (<>) and often come in pairs. The end tag, the second tag in the pair, looks like the begin tag, the first tag in the pair, but with an additional / in front.

The format for an HTML document is

```
<html>
<head>
<title> text that is to appear as title </title>
</head>
<body>
        text that is to appear as body on the page
</body
</html>
```

Self Check Exercise

4) Write an SQL statement to display the books issued from the library for a member named Amit.

Note: i) Write your answer in the space given below.

ii) Check your answer with the answers given at the end of this unit.

.....

.....

.....

.....

.....

.....

.....

.....

3.5 GENERATIONS OF PROGRAMMING LANGUAGES

Programming can be done in any language such as C, C++, COBOL, etc., which largely depends according to need. But a computer executes programs only after they are represented internally in binary form (sequences of 1s and 0s). Programs written in any other language must be translated to the binary representation of the instructions before they can be executed by the computer.

Just as hardware is classified into generations based on technology, computer languages also have generations classification based on the level of interaction with the machine. Since 1945 over the years, five categories or generations of programming languages have evolved over the years which are as follows:

- First generation, 1945 – *Machine language*
- Second generation, mid 1950s – *Assembly language*
- Third generation, early 1960s – *High level languages*
- Fourth generation, early 1970s – *Very high-level languages*
- Fifth generation, early 1980s – *Natural languages*

Programs written for a computer may be in either one of the categories of languages.

3.5.1 Machine Language

This is a sequence of instructions written in the form of binary numbers consisting of 1s and 0s which the computer responds directly. The machine language was initially referred to as code, although now the term code is used more broadly to refer to any program text.

An instruction prepared in any machine language will have at least two parts. The first part is the Command or Operation, which tells the computer what functions, is to be performed. All computers have an operations code for each of its functions. The second part of the instruction is the operand. It tells the computer 'where to find' or 'store the data' that has to be manipulated. Machine language is considered to be the first generation language.

3.5.2 Assembly Language

When symbols (letters, digits or special characters) are used for the operation part, the address part and other parts of the instruction code, this representation is called an assembly language program. This is considered to be second-generation language. Machine and Assembly languages are referred to as low-level languages since the coding for a problem is at the individual instruction level. Each machine has got its own assembly language which is dependent upon the internal architecture of the processor. Both machine and assembly languages are machine dependent.

Since a computer can execute programs only in machine language, a translator or converter is needed if the program is written in any other language. An **assembler** is such a language translator which takes its input in the form of an assembly language program and produces machine language code in some memory location.

3.5.3 High-level Languages

High-level languages are also called procedural languages. Programming languages such as C, COBOL, FORTRAN and BASIC are high-level languages. As the time and cost for creating machine and assembly languages are high, the high level languages were developed.

A program written in a high level language should be translated into a form the machine can understand and this is done by a software (language translator) called **compiler** which takes the source code input and produces as output the machine language code of the machine on which it is to be executed. There is another kind of software which also does the translation. This is called an **interpreter.**

3.5.4 Fourth Generation Languages

Fourth Generation Language, referred to as 4GL is a high level language that requires significantly fewer instructions than a third generation language does to accomplish a particular task. Thus, the programmer should be able to write a program faster in 4GL than in third generation language. Most Fourth Generation Languages are non-procedural languages. The programmer does not have to give the details of procedure in a program, but instead specifies what is wanted.

For example, assume you need to display the details of a particular book (*Introduction to Computers)* from the catalogue. In a procedural language, the programmer would have to write a series of instructions in the following steps:

Step 1: Get a record from the catalogue

Step 2: If this is the record on *Introduction to Computers*, display the details

Step 3: If this is not the record for *Introduction to Computers*, go to Step 1

In a 4GL, however, the programmer would write a single instruction that says:

Get the details of *Introduction to Computers* from the Catalogue

Majority of fourth generation languages are used to get information from files and databases and display or print the information. These fourth generation languages contain a query language which is used to answer queries or questions with data from a database.

Structured Query Language (SQL) is a standard language and universally used. Like most other computer languages like C, C++, Pascal, or Java, you need a way to use it on your computer.

Most database applications use the SQL, as their language to communicate with the database, like add, delete, modify the data.

3.5.5 Natural Languages

Natural languages are of two types. The first are ordinary human languages like English. The second are programming languages that use human language to give people more natural interaction with computers. Natural languages allow questions or commands to be framed in a more conversational way. Natural languages are part of the field of study known as *artificial intelligence.* Artificial intelligence is a group of related technologies that attempt to develop machines capable of emulating human-like qualities, such as learning, reasoning, communicating, seeing, and hearing.

Self Check Exercise

5) Name some third generation languages.

Note: i) Write your answer in the space given below.

ii) Check your answer with the answers given at the end of this unit.

.....

.....

.....

.....

.....

.....

.....

.....

3.6 SUMMARY

In this Unit, we have dealt with programming and programming languages. The focuses of attention in this Unit are the following:

- Concept of programming and programming languages;
- Steps in programming;
- Machine, assembly and high level languages;
- Common programming languages and uses.

3.7 ANSWERS TO SELF CHECK EXCERCISES

1) Program.

2) Algorithm consists of a set of explicit and unambiguous finite steps which, when carried out for a given set of initial conditions, produce the corresponding output and terminate in a finite time. A pictorial representation of the algorithm using standard symbols is a flow chart.

3) The six steps in programming are (1) Program Specification, (2) Program Design, (3) Program Code, (4) Program Test, (5) Program Documentation and (6) Program Maintenance. Coding is one of the steps in programming that involves the actual writing or keying in the set of instructions in a programming language.

4)
```
SELECT BOOK_ISSUED
FROM CIRCULATION_DATA
WHERE BORROWER_NAME = "Amit"
```

5) Some third generation languages are C, C++, BASIC and FORTRAN.

3.8 KEYWORDS

Application program : Program that has been developed to solve a particular problem, perform useful work on general-purpose tasks, or provide entertainment

Assembly language : It is a low level programming language that allows a computer user to write a program using abbreviations or more easily remembered words instead of numbers.

Compiler : A language translator that converts the entire program of a high-level language into machine language before the computer executes the program.

High level language : Also known as the procedural language it resembles human languages such as English.

Interpreter : A language translator that converts each procedural language statement into machine language and executes it immediately.

Machine language : A binary type low-level language consisting of 1s and 0s that the computer can run directly.

System program : Program that helps the computer to perform essential operating tasks and enables the application program to run.

3.9 REFERENCES AND FURTHER READING

Droomy, R.G. (2004). *How to solve it by computer*. New Delhi: Prentice Hall of India.

O'Leary, Timothy J. and O'Leary, Linda I. (2002). *Computing essentials 2002-2003*. International Edition. New York: McGraw-Hill Irwin.

Pratt, Terrence W. and Zelkowitz, Marvin V. (2003). *Programming languages: design and implementation*. 4th Edition. New Delhi: Pearson Education.

Schneider, G. Michael and Gersting, Judith L. (1998). *An introduction to computer science*. 2nd Edition. California: Brooks/Cole Publishing Company.

Williams, Brian K. and Sawyer, Stacey C. (2003). *Using information technology: a practical introduction to computers and communications*. 5th Edition. New Delhi: Tata McGraw-Hill Publishing.

ignou
THE PEOPLE'S
UNIVERSITY